IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 April 2024, accepted 12 May 2024, date of publication 16 May 2024, date of current version 6 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3402125

RESEARCH ARTICLE

A Universal Field-of-View Mask Segmentation Method on Retinal Images From Fundus Cameras

V. V. STAROVOITOV[1], NGUYEN NHU SON[2], YU. I. GOLUB[1], M. M. LUKASHEVICH[3], NGUYEN LONG GIANG[2], HOANG THI MINH CHAU[4], AND LE HOANG SON[5]

[1]United Institute of Informatics Problems, National Academy of Sciences of Belarus, 220072 Minsk, Belarus
[2]Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi 10000, Vietnam
[3]Department of Electronic Computer Machine, Belarusian State University of Informatics and Radioelectronics, 220013 Minsk, Belarus
[4]Department of Information Technology, University of Economics and Technology for Industries, Hanoi 10000, Vietnam
[5]VNU Information Technology Institute, Vietnam National University, Hanoi 10000, Vietnam

Corresponding author: Nguyen Nhu Son (nnson@ioit.ac.vn)

This work was supported in part by Vietnam Academy of Science and Technology under Grant QTBY01.09/22-23, and in part by Belarusian Republican Foundation for Fundamental Research under Project F22V-010.

ABSTRACT One of the first steps in the retinal image preprocessing is cropping the Field of View (FOV) area and scaling it into a template of a predefined size. Fundus cameras of different producers record digital images of the retina of various sizes, and the FOV area containing helpful information can be from 43 to 98% of the image area. For automated analysis of retinal images and detection of DR, it is necessary to segment the FOV region and cut it out from the image. This is important to preserve microaneurysms and small capillaries in the retinal image as much as possible, since neural network methods always reduce the original images to a predefined size. In this paper, we propose a universal method for FOV segmentation based on the ideas of histograms and thresholds. We compared 11 methods for segmenting FOV regions on the four most commonly used retinal image grayscale representations. In total, we compared 35 variants of segmentation and evaluated the obtained results by four functions: Jaccard index, Matthews correlation coefficient (MCC), accuracy and balanced accuracy. All options were tested on 7000 images from nine of the largest databases. The images were generated by 100 different fundus cameras. The following observations have been extracted through extensive comparative experiments namely: 1) segmentation of the FOV area should be performed on the grayscale image obtained from the red channel; 2) for more accurate segmentation, a logarithmic transformation should be applied to the grayscale image; 3) the FOV area mask can be segmented by a global threshold calculated by Otsu's method; 4) global thresholding based on analysis of histogram peaks does not provide advantages over binarization by Otsu's method applied to the logarithmic transformation of the image.

INDEX TERMS Fundus image, field of view (FOV) mask, FOV segmentation, dataset, image representation.

I. INTRODUCTION

Currently, more and more attention is being paid to screening in medicine. Screening is a set of diagnostic procedures aimed at identifying diseases in clinically asymptomatic individuals. Many diseases begin asymptomatically, and some of them cannot be cured. For example, most patients with diabetes asymptomatically develop diabetic retinopathy, which can lead to blindness. This disease can only be diagnosed by analyzing the changes occurring in the patient's retina.

The International Diabetes Federation has announced that 537 million people worldwide have diabetes, and many develop retinopathy, which can lead to blindness [1]. The World Health Organization (WHO) predicts that India will

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

have the largest number of adults with diabetes in the world, reaching 80 million in 2030 [2]. Approximately 50% of people worldwide are undiagnosed with diabetes [3].

Diagnosing retinal diseases at the earliest stage and promptly treating them can lead to slower loss of a patient's vision. However, it is impossible to ensure regular examination of the retina of every resident by ophthalmologists, not only in developing countries with huge populations, but also in developed countries. The number of ophthalmologists worldwide exceeds 200,000, but there is a significant shortage of ophthalmologists in developing countries. This is especially true for low- and middle-income countries with huge populations such as Bangladesh and India. For example, according to the International Council of Ophthalmology [4], in Bangladesh there are only four ophthalmologists per million people, and in India there are eleven.

In recent years, telemedicine has been increasingly developed. Its use is especially effective in underdeveloped countries where there is a shortage of medical facilities and qualified doctors. Patients with diabetes mellitus are recommended to be screened for DR every 2.5 years [5], and patients with this disease - once a year. To take advantage of the benefits of telemedicine, it is important to be able to capture images of the patient's retina with a special fundus camera and transmit them to the ophthalmology center. This can be done by a nurse.

In the article [6], the authors described the technological diversity of fundus photography systems for retinal screening. Analysis of retinal images collected in large public databases, such as Kaggle [7], shows that the retinal surface is photographed using different fields of view (FOV), the optical systems used and digital image acquisition matrices differ. All this variety of digital data needs to be converted into a single view for processing by machine learning systems that help the ophthalmologist make better decisions. One of the first steps in the data preprocessing stage is to crop the area of the FOV ROI from the digital image and scale it into a template of a predefined size [9]. The variety of initial retinal images generated by cameras from different manufacturers requires a unified method for segmenting the FOV area. The FOV mask is often used to remove the detected blood vessels outside the fundus image ROI.

In this paper, we propose a universal method for FOV segmentation based on the ideas of histograms and thresholds. The goal of this paper was to compare the main approaches to FOV area segmentation and find the best one which did not depend on the fundus camera type and image size. We compare all 11 methods listed in Table 2 on the four most commonly used retinal image representations. The total number of segmentation variants was 35. FOV area segmentation performance was assessed by three functions: Jaccard index, Matthews correlation coefficient (MCC), and balanced accuracy. All algorithmic variants were tested on 7,000 images from eight largest databases. Test images were created by 100 different fundus cameras.

We received the extensive results and tried to present the most interesting of them in the form of images for visual comparison of segmentation results by different methods, and collected numerical estimates in the tables and box plots.

The remainder of the paper is organized as follows: Section II presents a short review of diversity of FOV area representation in retina images and main approaches for its segmentation. Section III describes our experimental basis, including evaluation functions. Section IV describes two types of experiments and discuss the obtained results in detail. Finally, Section V concludes the paper.

II. MATERIALS AND METHODS

A. DATA PREPARATION

Every camera has a *Field of View* (FOV), which is a part of the fundus that is visible through the camera lens. The visible surface of the retina is the inner surface of a spherical body. It is concave, and its image, which fell into the FOV, is projected onto the plane of the sensor matrix. In the matrix, the visible part of the retina is represented as a circle surrounded by a dark background. However, some camera manufacturers cut off the edges of the circle at the top and bottom, while others cut off the left and right edges (Fig. **1**).

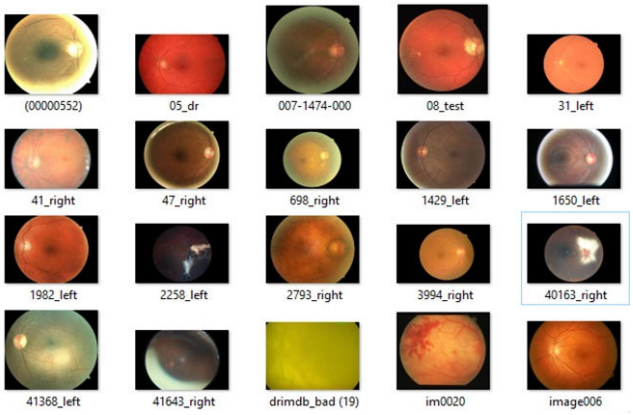

FIGURE 1. Variants of FOV area representation in sample images.

The variety of sizes of retinal images presented only in one publicly available Kaggle database [7] is shown in Table 1. The part of the image useful for analysis, which is the retina itself, ranges from 43% (image "3994_right" in Fig. **1**) to 98.5% (image "Drimdb_bad" in Fig. **1**) of the image matrix area.

One of the main approaches to analysis and classification of digital retinal images are based on artificial neural networks [8, 32-36]. Images of a fixed size from 224 × 224 to 512 × 512 pixels are fed to the input of the neural network (Table 1). The networks are trained to recognize signs of disease and classify the stage of the disease. It is desirable to minimize the presence of the background area in the input images. Therefore, at the first stage of retinal image analysis, the FOV region is segmented, then is cropping and scaled

TABLE 1. Kaggle fundus image sizes in pixels.

315×400	289×433
1184×1792	1664×2496
1880×2816	1920×2560
1050×1050	1226x1844
1944×2592	1957×2196
2000×3008	2048×3072
2056×2124	2056×3088
2136×3215	2304×3456
2336×3504	2448×3264
2592×3872	2592×3888
2848×4272	2848×4288
3168×4752	3264×4928
3456×5184	1216×1600

to the predefined size. We carried out a study on choosing the optimal size of the retinal image supplied to the network input, and came to the conclusion that it should be equal to 512 × 512 pixels. To maximize the preservation of the first signs of retinopathy (microaneurysms), it is better to cut out the square inscribed in the FOV area [9].

B. RELATED WORK

In [10], the authors analyzed over 100 papers and found that some authors evaluate the results of performing vessel segmentation classification only in the FOV region, while others use the entire image. There are no vessels outside the FOV. If the pixels are treated as true negatives, the accuracy of vessel segmentation and other scores will be different. In [10], the authors compared the results of one annotator in the FOV masks only with those of another annotator all image pixels. In the FOV mask, they got Accuracy (abbreviation below ACC) = 0.9473, Sensitivity = 0.7760 and Specificity = 0.9725, while using all pixels all scores were higher: ACC = 0.9636, Sensitivity = 0.7756 and Specificity = 0.9818. The difference in the ACC and Specificity values was more than 1.5% and 1%, respectively. This highlights the importance of accurate segmentation of the FOV region.

FOV segmentation methods can be divided into two groups according to the following concepts: i) presentation of initial data for segmentation (Red, Green, L from Lab, V from HSV, Grayscale, log intensity transformation); ii) Method of the threshold calculation (fixed, by Otsu binarization [11], by other binarization method, based on image statistics). Many binarization methods are described and compared in the review [12], but recommendations are given mainly for images representing scanned texts. Nevertheless, the Kittler & Illingworth method [13] was named one of the best. Our experiments have shown that the threshold value calculated by Kittler usually is less than the threshold obtained by Otsu method for the same image, but vice versa in the logarithmic intensity representation. If the valley separating the background and the FOV area on the intensity histogram is wide, results of binarization by both methods practically coincide; otherwise, the results will be different [14].

One popular approach of FOV area segmentation is to use the green channel for histogram calculation. Binarization is performed on the first non-zero value of the histogram [15]. Santhakumar et al. [16] proposed a method for automatic generation of FOV mask. The proposed method uses a logarithm of the green channel with subsequent application of the Otsu binarization algorithm and morphological closing.

Another possible approach is to convert a color image into a grayscale image. For a grayscale image, the histogram of brightness is calculated and the binarization threshold was set as maximum of the first order derivative [17]. In [18] the FOV mask is created to obtain the boundary of the eye fundus in the acquired color image by clipping the effective area taken by the camera. The procedure includes several stages: red channel extraction, green channel extraction, combination of green and red channels, applying median filtering and Otsu binarization. Upadhyay et al. [19] presented a pixel-wise weighted combination of all the three channels of RGB color space. These weights are optimized by minimizing the mean square error, using ground truth of FOV masks from publicly available fundus image databases. Such combined image is used for generation of the FOV mask image by using a single-valued intensity threshold. This threshold value is calculated using the Otsu method. Another solution is to choose the threshold as the mean value of the green channel of the image divided by 2. This approach was presented in [20].

C. THE PROPOSED FOV MASK SEGMENTATION ALGORITHM

Our FOV mask segmentation algorithm:

Input: a color retina image, and the threshold T.

Output: a segmented image.

Step 1. Upload a color retina image. Extract the red channel R.

Step 2. Calculate the minimum R_{min}, average R_{me} and median R_{med} values in R.

Step 3. Change R_{min} and R_{med}:

If $R_{med} < T_1$, $R_{min} = R_{med}$ (T_1 is a predefined constant);

If $R_{me} < T_1$, $R_{me} = 2 \times R_{me} - R_{med}$ and $R_{min} = R_{med}$.

Step 4. Plot the histogram H of 256 bins.

Step 5. Calculate the threshold $T = \min(H(R_{min})+\tilde{n}, R_{me}))$, where $\tilde{n}$ is an empirically chosen constant.

Step 6. Segment the FOV mask by binarization of the red channel R with the threshold T.

Note. $T_1 = 30$ and $c = 3$ were chosen experimentally.

The rest FOV segmentation algorithms and their short descriptions are collected in Table 2. All of them were tested on the red, green, L components (from Lab) and grayscale image representations. Final FOV segmentation was obtained by application of the global thresholds calculated by every method on four retinal image representation.

It should be noted that with uneven illumination of the retina, the FOV area cannot be correctly segmented. This means that it is impossible to calculate the correct value of the optimal global threshold for retina image binarization. An example is shown in Fig. 2. The original image from the STARE database is shown, and the correlation plot of

TABLE 2. Abbreviation and description of the compared algorithms.

No.	Abbreviation	Description	Ref.
1	Our method	Histogram based (see below)	
2	Otsu	Direct Otsu binarization implementation: T=Otsu(X)	[11]
3	Kittler	Direct Kittler binarization implementation: T=Kittler(X)	[13]
4	Otsu(Log)	I=log2(X+k), T=Otsu(I)	
5	Kit(Log)	I=log2(X+k), T=Kittler(I)	
6	Giancardo	Green channel, histogram(G), 1st order derivative, T=1stnonzero value	[15]
7	Santhakumar	I=log2(Green+k), T=Otsu(I)	[16]
8	Guo	RGB->Gray, calculate histogram, T=max of the 1st order derivative	[17]
9	Lara	I=Red/(Green+1), $I2$=medfilrt(I), T=Otsu($I2$)	[18]
10	Upad	RGB->Gray, T=Otsu(Gray)	[19]
11	Raut	T=mean(Green)/2	[20]

Note: X means a grayscale representation of RGB.

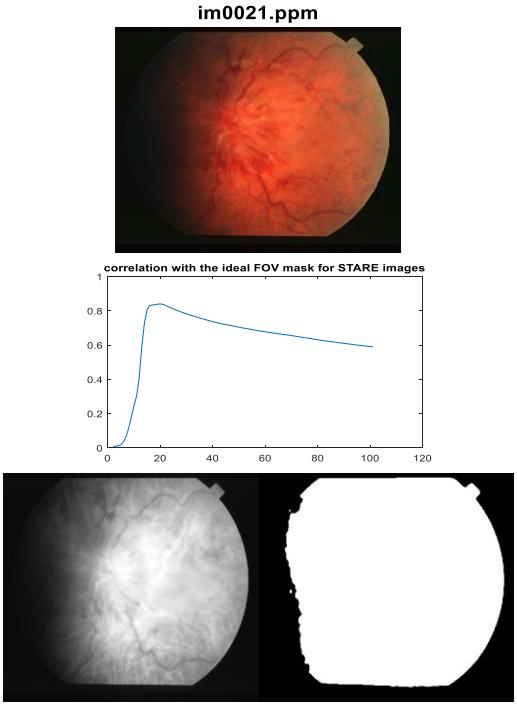

FIGURE 2. Image "im0021" from the STARE database, for which the FOV area cannot be accurately segmented.

the red channel segmentation with the reference mask when the threshold changes from 1 to 100. In the bottom, the red channel is shown, and the best result of its binarization by Kittler's method, the threshold is 20.

There are some problems with FOV area binarization based on histogram. Binarization algorithms work well if the histogram has two prominent peaks and a clear valley between them. These conditions are not always satisfied when binarizing different representations of retinal images. Otsu's and Kittler's algorithms missed the first valley in the red channel histogram of image "03a7f4a5786f" shown in the center of Fig. 3, but correctly detected the valley in the histogram of the logarithmic representation of the red channel (see Fig. 4).

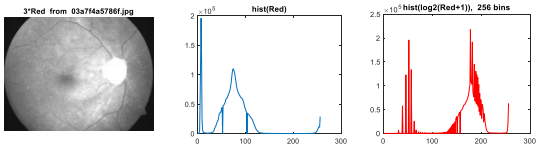

FIGURE 3. Red channel multiplied by 3, red channel histogram and histogram of logarithmic red.

In Figs. 4 and 6, the images are arranged in the following order, left to right: i) First line: initial grayscale image, template mask, Our method, Otsu, Kittler; ii) Second line: Otsu(log), Kittler(log), Santhakumar, Giancardo, Guo; iii) Third line: Lara, Upad, Raut.

In Fig. 4 one can see that the algorithms by Otsu and Kittler detected the optical disk in red channel, but not the FOV area. Therefore, in our algorithm using the histogram, additional details in Steps 2 and 3 are added for more correct FOV area binarization (see our mask as the third image in the middle of the top line in Fig. 4).

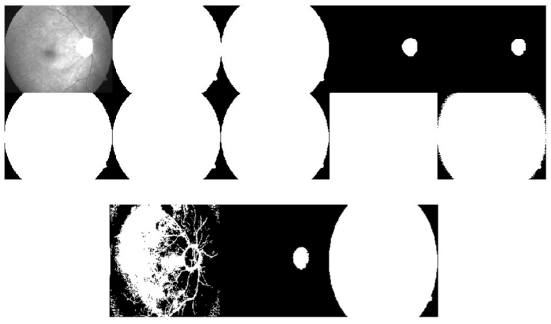

FIGURE 4. The results of image "03a7f4a5786f" binarization by 11 studied algorithms.

III. EXPERIMENTAL SETUP

A. DATABASES

Many papers connected to the FOV mask generation contain experiments based on a couple databases containing not much small fundus images. For example, Santhakumar [16] used DIARETDB1 database (89 images) and MESSIDOR database (460 images); Rodriguez [18] used DIARETDB1 database (89 images) and DIARETDB0 database (130 images).

We have used thousands of images from nine big open databases. Below we list the databases, and in brackets their subsets and how many images they contain. They are: Kaggle-2015 (2431 images from the subset train, category Mild) [7], STARE (398 images), DIARETDB (130 images in the DDB0 subset), DRIMDB (69 images from the Bad subset), DMED (169 images), MESSIDOR (515 images), RFMiD (640 images from the Test subset), DDR (2733 from

the Valid subset), HRF (45 images with individual FOV masks provided). To register these images, several dozens of cameras with different characteristics from different producers were used. The quality of the images is also different in different databases.

Due to the large amount of data, below we present the results of experiments on the four largest databases: Kaggle-2015, DDR, RFMiD, and STARE.

B. FUNCTIONS FOR RESULT EVALUATION

Fundus image binarization from one hand belongs to the binary classification tasks, from another – to two class segmentation. We utilized the most popular functions for classification and segmentation quality evaluation. They are accuracy (ACC), balanced accuracy (BalACC), F1, MCC, Jaccard index, Dice index. To evaluate our results, we used the ground truth segmentation masks and compare them with results obtained by different approaches.

Let for every fundus image the ideal FOV area is X set of pixels. Let Y is the set of pixels segmented by any method. From another side pixels of the ideal FOV area belong to the true positive set (TP) and the background pixels belong to the true negative set (TN). There errors when pixels are assigned to the wrong sets and we obtain false positive (FP) and false negative (FN) values. Therefore, for every method applied to every image we have a confusion matrix containing four elements (TP FP; FN TN).

We can now present scoring functions for both segmentation and classification approaches. In segmentation into two classes, we use background (B) and foreground (F), where subscript o means the result of segmentation, subscript t – ground truth. From classification point of view $TP = |F_o \cap F_t|$, $FN = |F_t| - |F_o \cap F_t|$ and so on. Here $|\ldots|$ means the number of the set elements. The formulas used in our paper are: Jaccard index, Dice index, Accuracy (ACC), balanced Accuracy (BalACC) and the normalized variant of the Matthews correlation coefficient (MCC) as presented below:

$$Jaccard = \frac{|F_o \cap F_t|}{|F_o \cup F_t|} = \frac{TP}{TP + FN + FP},$$

$$Dice = \frac{2 \times TP}{2 \times TP + FN + FP} = F1 = \frac{2 \times Jaccard}{1 + Jaccard},$$

$$ACC = \frac{|B_o \cap B_t| + |F_o \cap F_t|}{|B_t| + |F_t|} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$BalACC = \frac{1}{2}\left(\frac{B_o \cap B_t}{B_t} + \frac{F_o \cap F_t}{F_t}\right) = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right),$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}},$$

$$MCCnorm = \frac{(MCC + 1)}{2}.$$

Chicco and Jurman claimed that MCC gives better scores in binary classification than ACC and F1 scores. Setiawan also declared that MCC is more informative in evaluation of image segmentation results than ACC, sensitivity, specificity, Dice coefficient, and Jaccard index. Considering these studies and to reduce the size of the paper, we have excluded some evaluation functions and left only four in our study: MCC-norm instead of MCC to make its range is [0; 1] similar to other functions.

C. SOME PROPERTIES OF THE SCORE FUNCTIONS

We use the term FOV ratio as one of the parameters. It equals to the ideal FOV area divided on the number of pixels in the image. Jaccard and Dice indexes are very related and they do not take into account the TN value, the number of correctly segmented background pixels. That is why if a segmented FOV area is a full image, the Jaccard index is equal to the FOV ratio. If a segmented FOV mask is black (no detected FOV pixels), the Jaccard index is equal to zero.

ACC takes into account all four parameters of the confusion matrix, but it is known by its accuracy paradox. In our case that means the following: if the FOV ratio is very high (the highest is 0.98), ACC = FOV ratio when the segmented FOV mask will be a whole image. If the segmented FOV mask is black, ACC = 1 – FOV ratio.

Balanced ACC is good for a disbalanced classification, but in our case when the FOV ratio is very high the background pixels will dominate. For example, let the FOV ratio is 0.98, all the background pixels are classified correctly, but only 60% the FOV pixels are classified correctly, then BalACC = 0.5 + 0.254 = 0.754.

Since the Dice and Jaccard indices are essentially related, we used only the latter for economy.

IV. EXPERIMENTAL RESULTS

To obtain reference FOV masks, we segmented good quality retinal images from every database. We assumed that an arbitrary fundus camera generates digital images of a fixed size according to the built-in CCD matrix and the FOV area in this matrix is located according to the producer's own rules (see Fig. 1). Images from different cameras have different sizes. We found only one exception in the huge kaggle-2015 database (there were two different mappings of the FOV area on two matrices of the same size, Fig.5). Therefore, we used the size of the retinal image as an identifier for the reference FOV mask for all images of the same size. For example, if the image has dimensions of 1500 × 1152 pixels (database DDB0), the corresponding mask image is called "1500 × 1152mask" and is written in *tif* format. Every mask contains zeros for background and ones for the FOV area. Note, real FOV masks for fundus images of the same size never coincide on 100% due to different eye position during the registration and the anatomical retina surface variability, but for comparative segmentation quality analysis one good reference is enough.

Experiment 1: We manually select several bad quality images (for images of good quality no problem to segment a mask by any method) from the datasets described above.

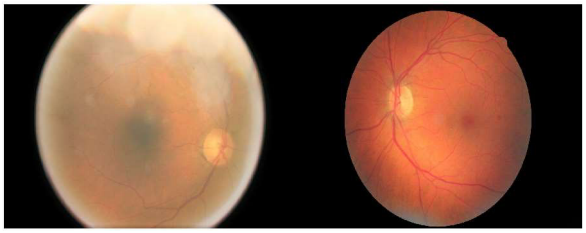

FIGURE 5. Images 24341_left and 17009_right from Kaggle-2015 dataset with the same size image size 2592 × 3888 pixels.

We apply the algorithms described above (except three last) to red R, green G, L components (from Lab) and to gray fundus image representation. The last three are applied to transformations based on combinations of red and green data. We compare 35 obtained FOV masks generated for every original image and calculate four score values: Jaccard index, Accuracy (ACC), balanced accuracy(BalACC) and the normalized Matthews correlation coefficient(MCCnorm). They all have the same value range of [0, 1]. We display all the masks and the corresponding scores. The goal of this experiment was to select the better fundus image representation and the best approach for FOV mask segmentation.

Experiment 2: We apply the same procedure but to all images presented in every dataset described above and calculate some statistics for every dataset concerning to the FOV mask quality: mean, median, minimum, maximum, standard deviation (std) and mean-4×std values. The larger the first four and sixth parameters and the smaller the std, the more accurately the FOV area is segmented. The most important score in our experiments is (mean-4×std) value. Then we select the better value of the four quality scores for every (from four) image representation and try to define the best approach for FOV area segmentation.

A. DISCUSSION OF EXPERIMENT 1 RESULTS

We found some images which may be difficult for FOV area segmentation. The most difficult images are presented in STARE and Kaggle databases.

Let us consider the generation of the FOV masks by 11 algorithms on 4 image representations and evaluate the results of segmentation by four score metrics using like an example ''0008'' image from the STARE database. In Fig.6 (a-c) results of segmentation by eight first algorithms from Table 2 on four representations of the original image are presented: Red, Green, L (from Lab) and Gray. The original representation is shown at the top left, the reference mask is shown to the right of it, then the masks obtained by our algorithm, Otsu, Kittler; next row: Otsu(log), Kittler(log), Santhakumar, Giancardo, Guo. Bottom line of images: Lara, Upadhyay, Raut.

Four types of assessments comparing the binarization results of image ''im0008'' with a reference FOV mask called ''605 × 700 mask'' and having FOV_ratio = 0.7362 are collected in Table 3. The highest scores are in bold. The best segmentation results for this image was obtained on the red channel. The best scores are highlighted in green and the worst in yellow in Table 3 (and in Table 4 below). They are shown by our algorithm and by the Kittler's algorithm.

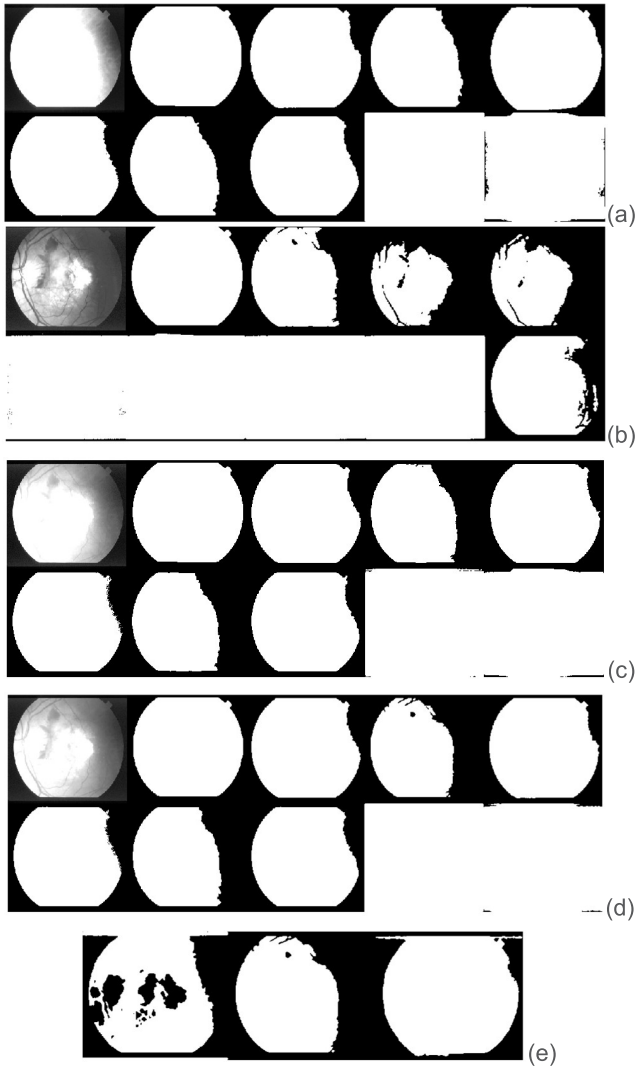

FIGURE 6. FOV mask segmented in R (a), G (b), L (c) and gray (d-e) representations of image ''im0008'' from STARE database.

In Green representation, four algorithms (Otsu(Log), Kit(Log), Giancardo, Santhakumar) indicated in Table 2 produced white squares. They did not segment the FOV mask. On the other three representations, only two algorithms produced similar white squares. The G channel gives a darker image representation with less contrast on the right. The Jaccard index does not evaluate segmentation errors when the mask pixels exist beyond the boundaries of the reference. For example, when evaluating segmentation by the Giancardo algorithm on the Gray representation, the index is equal to 0.7373, i.e. almost equal to the FOV ratio of this image. If the white square will be used instead of a real mask and the FOV ratio is 0.9612, the segmentation result (according to the Jaccard index) is very good. Accuracy of the same algorithms is approximately equal to 0.67, i.e. this score notes that the background is poorly segmented.

TABLE 3. Evaluation scores of the compared FOV binarization algorithms for im0008.ppm.

Red	Jaccard	BalACC	MCC norm	Accuracy
Our	**0.9510**	0.9475	**0.9537**	**0.9634**
Otsu	0.7725	0.8313	0.8018	0.8382
Kittler	**0.9538**	**0.9649**	**0.9624**	0.9543
Otsu (Log)	0.9248	0.9440	0.9153	0.9334
Kittler (Log)	0.7423	0.8091	0.7866	0.8232
Gian	0.7362	0.7363	0.8681	0.5089
Santhakumar	0.9247	0.9440	0.9152	0.9334
Guo	0.7650	0.7739	0.8808	0.6652
Green	Jaccard	BalACC	MCC norm	Accuracy
Our	0.7265	0.7782	0.8143	0.7971
Otsu	0.4880	0.6218	0.7016	0.7200
Kittler	0.5306	0.6530	0.7117	0.7349
Otsu (Log)	0.7674	0.7784	0.8392	0.6706
Kittler (Log)	0.7679	0.7781	0.8648	0.6726
Giancardo	0.7673	0.7780	0.8420	0.6702
Santhakumar	0.7673	0.7782	0.8382	0.6703
Guo	0.8464	0.8859	0.8475	0.8796
L	Jaccard	BalACC	MCC norm	Accuracy
Our	0.9445	0.9369	0.9493	**0.9587**
Otsu	0.7259	0.7969	0.7788	0.8152
Kittler	0.9484	**0.9616**	0.9415	**0.9525**
Otsu (Log)	0.9359	0.9523	0.9272	0.9423
Kittler (Log)	0.7404	0.8076	0.7855	0.8220
Giancardo	0.7400	0.7414	0.8700	0.5604
Santhakumar	0.9285	0.9468	0.9191	0.9364
Guo	0.7557	0.7621	0.8739	0.6359
Gray	Jaccard	BalACC	MCC norm	Accuracy
Our	0.9441	0.9364	0.9490	**0.9584**
Otsu	0.6894	0.7700	0.7631	0.7983
Kittler	0.9545	**0.9661**	0.9495	**0.9575**
Otsu (Log)	0.9260	0.9449	0.9165	0.9344
Kittler (Log)	0.7608	0.8226	0.7954	0.8318
Giancardo	0.7362	0.7363	0.8681	0.5089
Santhakumar	0.9211	0.9413	0.9114	0.9307
Guo	0.7453	0.7484	0.8726	0.5931
Last 3 algorithms	Jaccard	BalACC	MCC norm	Accuracy
Lara	0.7004	0.7646	0.7263	0.7516
Upadhyay	0.6900	0.7704	0.7633	0.7986
Raut	0.9511	**0.9630**	0.9557	**0.9520**

TABLE 4. Evaluation scores of the compared FOV binarization algorithms for drimdb_bad(17) image.

Red	Jaccard	BalACC	MCC norm	Accuracy
Our	**0.9679**	**0.9840**	0.7088	**0.9682**
Otsu	0.2768	0.3048	0.5262	0.5601
Kittler	0.1105	0.1450	0.5217	0.5346
Otsu (Log)	0.7303	0.7407	0.5646	0.6529
Kittler (Log)	0.1317	0.1654	0.5222	0.5382
Giancardo	0.9671	0.9673	**0.9836**	0.6959
Santhakumar	0.7259	0.7364	0.5638	0.6516
Guo	0.9461	0.9481	0.7136	0.8167
Lara	0.3140	0.3195	0.4612	0.4046
Upadhyay	0.2816	0.3094	0.5264	0.5608
Raut	0.6013	0.6167	0.5455	0.6163

In Red image representation, the Kittler's segmentation algorithm has the best scores: Jaccard = 0.9538, BalACC = 0.9649, MCCnorm = 0.9624, only ACC = 0.9543 gave the second result. Visual analysis of the results in Fig.7 this confirms. However, this algorithm is unstable when the retina is illuminated unevenly.

FIGURE 7. Drimdb_bad (17) from Drimdb database.

If a binarization algorithm applied to an image from the STARE database (the size of all mages is 605 × 700 pixels) produces a black square, which is possible for a very dark image representation, then the estimates of the results of the segmentation will be as follows: Jaccard = 0.0, ACC = 0.2638, BalACC = 0.1319 and MCCnorm =0.500 (because MCC = 0.0). On an image of a different size, for example, 3264 × 4928 pixels, when getting a black square Jaccard = MCC = 0.0, ACC = 0.5029, BalACC = 0.2515. This means that ACC and BalACC are responsive to image size, but Jaccard and MCC are not. At the same time, MCC partially takes into account how correctly the background is segmented, while Jaccard does not.

Consider binarization of a bad quality image like "drimdb_bad (17)" from Drimdb database presented in Fig.7. Its FOV ratio = 0. 9612. The results of binarization of this image by the first 8 algorithms from Table 2 are presented in Fig. 8. Our method showed the best results (top central mask). If, as a result of binarization, the entire image is converted into a black rectangle, this will be a very bad result, but we will get the following estimates: Jaccard = 0.0, ACC = 0.0388, BalACC = 0.0194, MCCnorm = 0.5000. This means that no one estimate can be true. The four types of binarization scores for the 11 studied algorithms on the four grayscale image representations of this image are summarized in Table 4. Our algorithm gave the 3 best scores. Our mask is in the center of the first line in Fig.8.

FIGURE 8. Binarization of Red channel of image drimdb_bad (17) from Drimdb database.

Based on the results of the first experiment, five algorithms were selected that segment poor quality images better than others. They are: Our, Kittler(Log), Otsu(Log), Santhakumar and Raut. A more detailed comparison of these algorithms was carried out in the second experiment on sets of images from the databases listed above. The statistics of the four evaluation functions for each of the four image presentations

were calculated separately, since the sets of fundus cameras recording images in different databases did not overlap.

The DIARET_DDB0 database contains 130 color retina images and for every image a binary mask of the FOV area is presented in the database. However, many of these masks are poorly segmented, see Fig.9. They cannot be used in any research.

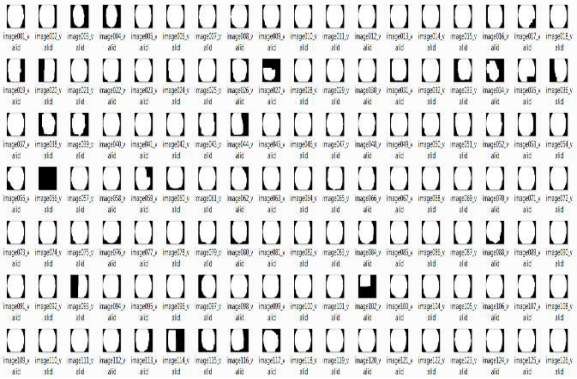

FIGURE 9. FOV masks provided in the DIARET_DDB0 database.

B. DISCUSSION OF EXPERIMENT 2 RESULTS

1) KAGGLE-2015 DATABASE

We used images from the Train directory, Mild subcategory. It contains 2443 images captured by 30 different types of cameras. We have generated 30 reference FOV masks of different size.

We have empirically determined that S = (mean – 4 × std) will be the main parameter for evaluating FOV segmentation quality in the tables below. The value S = 1 corresponds to ideal mask segmentation. The rest parameters are secondary. Jaccard index values were equally or very close to ACC values, that is why we present only Jaccard indexes in the text below. MCCnorm and balanced ACC (BalACC) give a bit different evaluation scores. We selected the maximal values of these three metrics for four types of retinal image representation (Red, Green, L and Gray) and ranked the evaluated algorithms. The ranks and types of presentations are indicated in the last column of the tables below. We then try to conclude which algorithm is better and in which image representation for every database described above. In parallel, we try to define the most objective metric for the FOV area segmentation.

For images from kaggle-2015 database we may say the following: 1) the R channel is the best image representation, 2) result-ranked algorithms are: Otsu(Log), Santhakumar, Our and Kittler(Log). Scores of all segmentation algorithms are very close. One may use a simpler algorithm. All scores are presented in Tables 5-7 and Fig.10. The best scores in the tables are highlighted in bold.

TABLE 5. Jaccard scores for images from Kaggle-2015 in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	**R=0.9708** G=0.9666 L=0.9679 Gr=0.9689	**R=0.9780** G=0.9733 L=0.9762 Gr=0.9758	R=0.0387 G=0.0452 L=0.0505 Gr=0.0411	R=0.5951 G=0.3336 L=0.0348 Gr=0.5992	R=0.8158 G=0.7860 L=0.7657 Gr=0.8045	R-3
Kittler(Log)	R=0.9701 G=0.9691 L=0.9701 Gr=0.9706	R=0.9775 G=0.9757 L=0.9776 Gr=0.9771	R=0.0391 G=0.0418 L=0.0422 Gr=0.0389	R=0.6508 G=0.4656 L=0.6064 Gr=0.6601	R=0.8136 G=0.8018 L=0.8014 Gr=0.8149	R-4
Otsu(Log)	R=0.9704 G=0.9670 L=0.9688 Gr=0.9687	R=.0.9767 G=0.9710 L=0.9745 Gr=0.9726	**R=0.0375** G=0.0423 L=0.0411 Gr=0.0387	**R=0.6706** G=0.4109 L=0.6732 Gr=0.6730	**R=0.8205** G=0.7978 L=0.8044 Gr=0.8141	R-1
Santhakumar	R=0.9704 G=0.9673 L=0.9683 Gr=0.9689	R=0.9771 G=0.9719 L=0.9746 Gr=0.9734	R=0.0384 G=0.0437 L=0.0436 Gr=0.0394	R=0.5951 G=0.2894 L=0.5088 Gr=0.5992	R=0.8167 G=0.7924 L=0.7940 Gr=0.8111	R-2
Raut	0.9684	0.9749	0.0404	0.4656	0.8067	R-5

TABLE 6. MCCnorm scores for images from Kaggle-2015 in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9787 G=0.9750 L=0.9768 Gr=0.9769	R=0.9847 G=0.9840 L=0.9843 Gr=0.9844	R=0.0274 G=0.0325 L=0.0327 Gr=0.0303	R=0.7654 G=0.7249 L=0.5538 Gr=0.7297	R=0.8690 G=0.8449 L=0.8462 Gr=0.8556	R-4
Kittler(Log)	R=0.9786 G=0.9776 L=0.9786 **Gr=0.9788**	R=0.9852 G=0.9845 L=0.9853 **Gr=0.9854**	R=0.0270 G=0.0289 L=0.0290 Gr=0.0275	R0.7766 G=0.7766 L=0.7770 Gr=0.7606	R=0.8704 G=0.8621 L=0.8627 Gr=0.8690	R-2
Otsu(Log)	R=0.9782 G=0.9755 L=0.9771 Gr=0.9767	R=0.9850 G=0.9832 L=0.9839 Gr=0.9836	R=0.0270 G=0.0296 L=0.0290 Gr=0.0282	**R=0.7777** G=0.7638 L=0.7736 Gr=0.7776	R=0.8702 G=0.8572 L=0.8609 Gr=0.8640	R-3
Santhakumar	R=0.9783 G=0.9759 L=0.9770 Gr=0.9770	R=0.9850 G=0.9833 L=0.9839 Gr=0.9838	R=0.0273 G=0.0297 L=0.0299 Gr=0.0283	R=0.7654 G=0.7149 L=0.7511 Gr=0.7742	R=0.8690 G=0.8573 L=0.8574 Gr=0.8638	R-4
Raut	0.9774	0.9835	**0.0261**	0.7592	**0.8728**	R-1

TABLE 7. BalACC scores for images from Kaggle-2015 in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9829 G=0.9814 L=0.9818 Gr=0.9824	R=0.9859 G=0.9852 L=0.9856 Gr=0.9857	R=0.0214 G=0.0233 L=0.0250 Gr=0.0221	R=0.7719 G=0.7751 L=0.6667 Gr=0.7731	R=0.8974 G=0.8883 L=0.8816 Gr=0.8942	R-3
Kittler(Log)	R=0.9821 G=0.9820 L=0.9823 Gr=0.9827	**R=0.9871** G=0.9859 L=0.9865 Gr=0.9867	R=0.0227 G=0.0222 L=0.0231 Gr=0.0214	R=0.8054 **G=0.8171** L=0.7754 Gr=0.7801	R=0.8915 G=0.8933 L=0.8899 Gr=0.8970	Gr-4
Otsu(Log)	**R=0.9831** G=0.9817 L=0.9821 Gr=0.9825	R=0.9864 G=0.9854 L=0.9858 Gr=0.9858	**R=0.0203** G=0.0220 L=0.0226 Gr=0.0207	R=0.7879 G=0.7907 L=0.7907 Gr=0.8107	**R=0.9018** G=0.8935 L=0.8918 Gr 0.8998	**R-1**
Santhakumar	R=0.9830 G=0.9817 L=0.9817 Gr=0.9825	R=0.9864 G=0.9856 L=0.9857 Gr=0.9858	R=0.0211 G=0.0228 L=0.0242 Gr=0.0214	R=0.7719 G=0.7751 L=0.7478 Gr=0.7731	R=0.8985 G=0.8906 L=0.8850 Gr=0.8967	R-2
Raut	0.9797	0.9850	0.0256	0.7534	0.8772	R-5

2) DDR DATABASE

We used 2733 images from the Valid directory. We have generated 55 FOV masks of different size. The best results on this base were shown by the Otsu(log), Santhakumar and Our algorithms on the R channel by all three measures. On green, all the algorithms had the worst results. All score values are collected in Tables 8-10 and Fig.11.

3) RFMID DATABASE

We used 1920 images from the Train directory. We have generated 3 FOV masks of different size. Results are presented

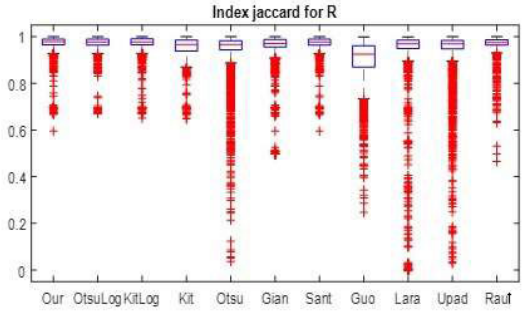

FIGURE 10. Box-plots of the Jaccard indexes for 11 algorithms applied to DDR dataset in red representation.

TABLE 8. Jaccard scores for images from DDR in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Rank
Our algorithm	R=0.9937 G=0.9907 L=0.9923 Gr=0.9923	R=0.9976 G=0.9976 L=0.9762 Gr=0.9758	R=0.0227 G=0.0416 L=0.0284 Gr=0.0276	R=0.4335 G=0.1074 L=0.4335 Gr=0.4335	**R=0.9030** G=0.8240 L=0.8787 Gr=0.8818	R-4
Kittler(Log)	R=0.9701 G=0.9891 L=0.9701 Gr=0.9706	R=0.9921 G=0.9972 L=0.9973 Gr=0.9771	R=0.0218 G=0.0385 L=0.0294 Gr=0.0292	R=0.5644 G=0.2913 L=0.5120 Gr=0.5070	**R=0.9048** G=0.8368 L=0.8745 Gr=0.8749	R-3
Otsu(Log)	R=0.9704 G= 0.9902 L=0.9922 Gr=0.9926	R=0.9975 G=0.9978 L=0.9745 Gr=0.9726	R=0.0199 G=0.0419 L=0.0287 Gr=0.0272	R=0.6242 G=0.2616 L=0.4832 Gr=0.5432	**R=0.9136** G=0.8224 L=0.8776 Gr=0.8839	R-1
Santhakumar	R=0.9704 G= 0.9902 L=0.9922 Gr=0.9927	R=0.9977 G=0.9978 L=0.9746 Gr=0.9734	R=0.0384 G=0.0451 L=0.0296 Gr=0.0273	R=0.5984 G=0.1074 L=0.4525 Gr=0.5353	**R=0.9124** **G**=0.8098 L=0.8737 Gr=0.8836	R-2
Raut	0.9863	0.9974	0.0515	0.2913	0.7801	R-5

TABLE 9. MCCnorm scores for images from DDR in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9964 G=0.9949 L=0.9818 Gr=0.9958	R=0.9988 G=0.9988 L=0.9988 Gr=0.9987	R=0.0121 G=0.0198 L=0.0148 Gr=0.0143	R=0.7167 G=0.7167 L=0.7167 Gr=0.7167	**R=0.9248** G=0.9158 L=0.9367 Gr=0.9387	R-3
Kittler (Log)	R=0.9955 G=0.9943 L=0.9954 Gr=0.9952	R=0.9980 G=0.9986 L=0.9987 Gr=0.9985	R=0.0154 G=0.0286 L=0.0212 Gr=0.0212	R=0.5102 G=0.3243 L=0.3360 Gr=0.3353	**R=0.9248** G=0.8797 L=0.9104 Gr=0.9105	R-3
Otsu (Log)	R=0.9961 G=0.9938 L=0.9953 Gr=0.9956	R=0.9988 G=0.9989 L=0.9990 Gr=0.9990	R=0.0115 G=0.0305 L=0.0176 Gr=0.0172	R=0.7702 G=0.3232 L=0.5278 Gr=0.5564	**R=0.9441** G=0.8717 L=0.9250 Gr=0.9267	R-1
Santhakumar	R=0.9962 G=0.9939 L=0.9953 Gr=0.9956	R=0.9989 G=0.9990 L=0.9990 Gr=0.9990	R=0.0116 G=0.0309 L=0.0180 Gr=0.0176	R=0.7610 G=0.3226 L=0.5192 Gr=0.5583	**R=0.9438** G=0.8701 L=0.9232 Gr=0.9253	R-2
Raut	0.9903	0.9986	0.0261	0.3623	0.8352	R-5

in Tables 11-13 and Fig. 12. Some scores indicate the best FOV segmentation on the L channel, but sum of three scores indicates that the R channel is the best for FOV segmentation. Otsu(Log) and Santhakumar are roughly the same, next is Our approach, while the Kittler(Log) algorithm is worse.

4) STARE DATABASE

There are 402 images in the database captured by 8 different cameras, and we created 8 reference FOV masks. This is a very difficult database for FOV segmentation. In Fig. 3 we present the first 90 images from this database. Results are presented in Tables 14-16 and Fig.14. The best image representation for FOV mask segmentation was the R channel. The best algorithms were Our, Otsu(Log), Santhakumar, while Kittler(Log) was worse.

The values of all evaluation functions differ, but are quite close, so we compare the different methods of FOV segmentation by ranking the most objective (in our opinion) estimate

TABLE 10. BalACC scores for images from DDR in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9964 G=0.9949 L=0.9818 Gr=0.9958	R=0.9988 G=0.9988 L=0.9988 Gr=0.9987	R=0.0121 G=0.0198 L=0.0148 Gr=0.0143	R=0.7167 G=0.7167 L=0.7167 Gr=0.7167	**R=0.9480** G=0.9158 L=0.9367 Gr=0.9387	R-3
Kittler (Log)	R=0.9955 G=0.9943 L=0.9954 Gr=0.9952	R=0.9980 G=0.9986 L=0.9987 Gr=0.9985	R=0.0154 G=0.0286 L=0.0212 Gr=0.0212	R=0.5102 G=0.3243 L=0.3360 Gr=0.3353	**R=0.9338** G=0.8797 L=0.9104 Gr=0.9105	R-4
Otsu (Log)	R=0.9961 G=0.9938 L=0.9953 Gr=0.9956	R=0.9988 G=0.9989 L=0.9990 Gr=0.9990	R=0.0115 G=0.0305 L=0.0176 Gr=0.0172	R=0.7702 G=0.3232 L=0.5278 Gr=0.5564	**R=0.9502** G=0.8717 L=0.9250 Gr=0.9267	**R-1**
Santhakumar	R=0.9962 G=0.9939 L=0.9953 Gr=0.9956	R=0.9989 G=0.9990 L=0.9990 Gr=0.9990	R=0.0116 G=0.0309 L=0.0180 Gr=0.0176	R=0.7610 G=0.3226 L=0.5192 Gr=0.5583	**R=0.9498** G=0.8701 L=0.9232 Gr=0.9253	R-2
Raut	0.9899	0.9987	0.0399	0.3178	0.8304	R-5

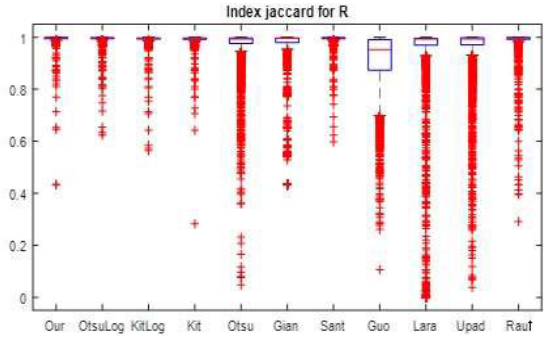

FIGURE 11. Box-plots of the Jaccard indexes for 11 algorithms tested on red channel of the DDR database images

TABLE 11. Jaccard scores for images from RFMiD in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Rank
Our algorithm	R=0.9937 G=0.9907 L=0.9923 Gr=0.9923	R=0.9976 G=0.9976 L=0.9762 Gr=0.9758	R=0.0227 G=0.0416 L=0.0284 Gr=0.0276	R=0.4335 G=0.1074 L=0.4335 Gr=0.4335	R=0.9030 G=0.8240 **L=0.9442** Gr=0.8818	L-4
Kittler(Log)	R=0.9701 G=0.9891 L=0.9701 Gr=0.9706	R=0.9921 G=0.9972 L=0.9973 Gr=0.9771	R=0.0218 G=0.0385 L=0.0294 Gr=0.0292	R=0.5644 G=0.2913 L=0.5120 Gr=0.5070	R=0.9048 G=0.8368 **L=0.9634** Gr=0.8749	L-1
Otsu(Log)	R=0.9704 G= 0.9902 L=0.9922 Gr=0.9926	R=0.9975 G=0.9978 L=0.9745 Gr=0.9726	R=0.0199 G=0.0419 L=0.0287 Gr=0.0272	R=0.6242 G=0.2616 L=0.4832 Gr=0.5432	R=0.9136 G=0.8224 L=0.8776 **Gr=0.9588**	**Gr-2**
Santhakumar	R=0.9704 G= 0.9902 L=0.9922 Gr=0.9927	R=0.9977 G=0.9978 L=0.9746 Gr=0.9734	R=0.0384 G=0.0451 L=0.0296 Gr=0.0273	R=0.5984 G=0.1074 L=0.4525 Gr=0.5353	R=0.9124 G=0.8098 L=0.8737 Gr=**0.9568**	Gr-3
Raut	0.9965	0.9986	0.0155	0.6930	0.9344	

TABLE 12. MCCnorm FOV segmentation estimates for images from RFMiD in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9976 G=0.9773 L=0.9969 Gr=0.9858	R=0.9993 G=0.9971 L=0.9992 Gr=0.9990	R=0.0093 G=0.0272 L=0.0084 Gr=0.0245	R=0.8782 G=0.8357 L=0.8753 Gr=0.8516	**R=0.9606** G=0.8686 L=0.9442 Gr=0.8879	R-4
Kittler (Log)	R=0.9930 G=0.9946 L=0.9984 Gr=0.9907	R=0.9990 G=0.9991 L=0.9992 Gr=0.9990	R=0.0315 G=0.0266 L=0.0067 Gr=0.0372	R=0.8143 G=0.8143 L=0.8778 Gr=0.8143	R=0.8671 G=0.8884 **L=0.9634** Gr=0.8421	L-3
Otsu (Log)	R=0.9974 G=0.9978 L=0.9980 Gr=0.9979	R=0.9984 G=0.9989 L=0.9990 Gr=0.9988	R=0.0079 G=0.0081 L=0.0073 Gr=0.0070	R=0.8461 L=0.8779 G=0.8783 Gr=0.8782	R=0.9660 G=0.9655 L=0.9556 **Gr=0.9698**	Gr-1
Santhakumar	R=0.9975 G=0.9978 L=0.9980 Gr=0.9979	R=0.9985 G=0.9990 L=0.9991 Gr=0.9989	R=0.0079 G=0.0082 L=0.0076 Gr=0.0073	R=0.8463 G=0.8746 L=0.8784 Gr=0.8783	**R=0.9659** G=0.9649 L=0.9529 Gr=0.9688	R-2
Raut	0.9979	0.9993	0.0098	0.7739	0.9586	

TABLE 13. BalACC scores for images from RFMiD in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	R=0.9964 G=0.9949 L=0.9818 Gr=0.9958	R=0.9988 G=0.9988 L=0.9988 Gr=0.9987	R=0.0121 G=0.0198 L=0.0148 Gr=0.0143	R=0.7167 G=0.7167 L=0.7167 Gr=0.7167	**R=0.9480** G=0.9158 L=0.9367 Gr=0.9387	R-3
Kittler (Log)	R=0.9955 G=0.9943 L=0.9954 Gr=0.9952	R=0.9980 G=0.9986 L=0.9987 Gr=0.9985	R=0.0154 G=0.0286 L=0.0212 Gr=0.0212	R=0.5102 G=0.3243 L=0.3360 Gr=0.3353	**R=0.9338** G=0.8797 L=0.9104 Gr=0.9105	R-4
Otsu (Log)	R=0.9961 G=0.9938 L=0.9953 Gr=0.9956	R=0.9988 G=0.9989 L=0.9990 Gr=0.9990	R=0.0115 G=0.0305 L=0.0176 Gr=0.0172	R=0.7702 G=0.3232 L=0.5278 Gr=0.5564	**R=0.9502** G=0.8717 L=0.9250 Gr=0.9267	R-1
Santhakumar	R=0.9962 G=0.9939 L=0.9953 Gr=0.9956	R=0.9989 G=0.9990 L=0.9990 Gr=0.9990	R=0.0116 G=0.0309 L=0.0180 Gr=0.0176	R=0.7610 G=0.3226 L=0.5192 Gr=0.5583	**R=0.9498** G=0.8701 L=0.9232 Gr=0.9253	R-2
Raut	0.9899	0.9987	0.0399	0.3178	0.8304	R-5

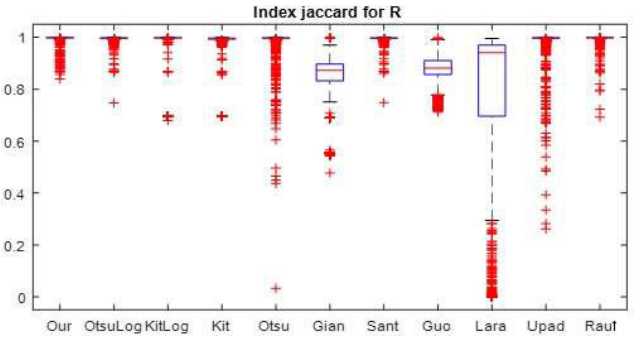

FIGURE 12. Box-plots for the Jaccard index for 11 algorithms tested on red channel of the RFMiD database images.

S =(mean – 4×std). It should be noted that, for example, the BalACC scores of three methods for segmenting images from RFMiD database (Table 13) are very close: 0.9502 by Otsu(Log), 0.9498 by Santhakumar, 0.9480 by Our method, but these methods have different ranks: 1, 2, 3. Summation of all S scores of the FOV area segmentation based on R channel gives 10.6779, 10.6631 and 10.6405 for the same score order. As one can see, the total absolute estimates are very close.

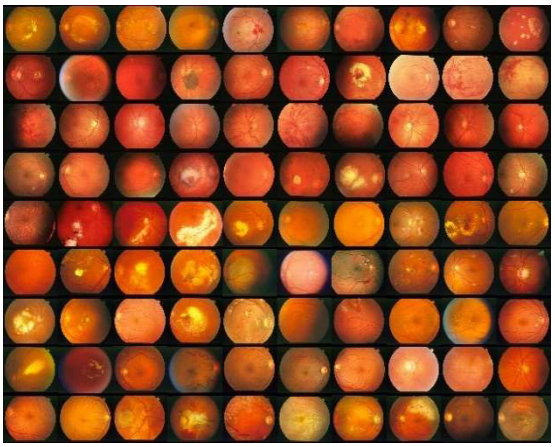

FIGURE 13. Ninety first images from the STARE database.

TABLE 14. Jaccard scores for images from STARE in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Rank
Our algorithm	**R=0.9407** G=0.8598 L=0.9075 Gr=0.9320	**R=0.9481** G=0.8928 L=0.9437 Gr=0.9459	**R=0.0354** G=0.0913 L=0.0765 Gr=0.0507	**R=0.7522** G=0.7265 L=0.7381 Gr=0.7377	**R=0.7989** G=0.4945 L=0.6013 Gr=0.7291	**R-1**
Kittler(Log)	R=0.9221 G=0.8998 L=0.9208 Gr=0.9228	R=0.9458 G=0.9374 L=0.9446 Gr=0.9455	R=0.0694 G=0.0936 L=0.0736 Gr=0.0721	R=0.5873 G=0.0625 L=0.2414 Gr=0.2515	R=0.6444 G=0.5254 L=0.6263 Gr=0.6344	R-4
Otsu(Log)	R=0.9387 G=0.8856 L=0.9356 Gr=0.9371	R=0.9465 G=0.9311 L=0.9466 Gr=0.9466	R=0.0378 G=0.0866 L=0.0442 Gr=0.0419	R=0.7409 G=0.4667 L=0.7394 Gr=0.7255	R=0.7876 G=0.5392 L=0.7587 Gr=0.7697	R-2
Santhakumar	R=0.9385 G=0.8872 L=0.9357 Gr=0.9368	R=0.9466 G=0.9336 L=0.9467 Gr=0.9466	R=0.0382 G=0.0869 L=0.0443 Gr=0.0422	R=0.7380 G=0.4788 L=0.7434 Gr=0.7265	R=0.7858 G=0.5395 L=0.7586 Gr=0.7681	R-3
Raut	0.9273	0.9404	0.0470	0.7512	0.7392	R-5

TABLE 15. MCCnorm scores for images from STARE in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	**R=0.9437** G=0.8195 L=0.8922 Gr=0.9306	**R=0.9501** G=0.8879 L=0.9461 Gr=0.9486	R=0.0324 G=0.1353 L=0.1095 Gr=0.0644	**R=0.6512** G=0.5998 L=0.6002 Gr=0.5991	R=0.8140 G=0.2783 L=0.4540 Gr=0.6729	**R-1**
Kittler(Log)	R=0.9335 G=0.8884 L=0.9269 Gr=0.9280	R=0.9486 G=0.9404 L=0.9478 Gr=0.9482	R=0.0467 G=0.1107 L=0.0645 Gr=0.0643	R=0.6493 G=0.5249 L=0.6377 Gr=0.6428	R=0.7466 G=0.4457 L=0.6688 Gr=0.6707	R-4
Otsu(Log)	R=0.9430 G=0.8620 L=0.9363 Gr=0.9398	R=0.9492 G=0.9328 L=0.9490 Gr=0.9491	**R=0.0322** G=0.1226 L=0.0520 Gr=0.0433	R=0.6496 G=0.6444 L=0.6533 Gr=0.6503	**R=0.8142** G=0.3717 L=0.7281 Gr=0.7667	R-2
Santhakumar	R=0.9429 G=0.8642 L=0.9365 Gr=0.9397	R=0.9492 G=0.9358 L=0.9493 Gr=0.9491	R=0.0323 G=0.1227 L=0.0519 Gr=0.0432	R=0.6505 G=0.6449 L=0.6536 Gr=0.6508	R=0.8138 G=0.3736 L=0.7290 Gr=0.7669	R-3
Raut	0.9255	0.9427	0.0572	0.6508	0.6969	R-5

Table 17 summarizes the ranked results of our experiments on the four largest databases listed above. Based on the results of our study, the following recommendations can be done: for the FOV area segmentation the logarithmic representation of red retina image channel should be used. The best segmentation algorithms are: Otsu(Log), Santhakumar and Our. They give very similar results.

TABLE 16. BalACC scores for images from STARE in red, green, L and gray representation.

	Mean	Median	Standard Deviation	Min	Mean-4Std	Best
Our algorithm	**R=0.9457** G=0.9112 L=0.9308 Gr=0.9418	**R=0.9517** G=0.9089 L=0.9473 Gr=0.9501	**R=0.0304** G=0.0476 L=0.0442 Gr=0.0343	R=0.7952 G=0.7737 L=0.7824 Gr=0.7699	**R=0.8240** G=0.7209 L=0.7539 Gr=0.8045	**R-1**
Kittler(Log)	R=0.9275 G=0.9237 L=0.9290 Gr=0.9309	R=0.9458 G=0.9402 L=0.9446 Gr=0.9459	R=0.0542 G=0.0511 L=0.0504 Gr=0.0487	**R=0.7209** G=0.5476 L=0.6586 Gr=0.6642	R=0.7108 G=0.7194 L=0.7275 Gr=0.7360	R-5
Otsu(Log)	R=0.9417 G=0.9194 L=0.9432 Gr=0.9425	R=0.9485 G=0.9390 L=0.9507 Gr=0.9500	R=0.0333 G=0.0514 L=0.0322 Gr=0.0336	R=0.7744 G=0.6659 L=0.7933 Gr=0.7818	R=0.8085 G=0.7139 L=0.8144 Gr=0.8080	R-2
Santhakumar	R=0.9414 G=0.9200 L=0.9430 Gr=0.9421	R=0.9484 G=0.9411 L=0.9509 Gr=0.9499	R=0.0336 G=0.0519 L=0.0324 Gr=0.0340	R=0.7747 G=0.6702 L=0.7944 Gr=0.7816	R=0.8070 G=0.7124 L=0.8134 Gr=0.8063	R-3
Raut	0.9391	0.9473	0.0346	0.7560	0.8007	R-4

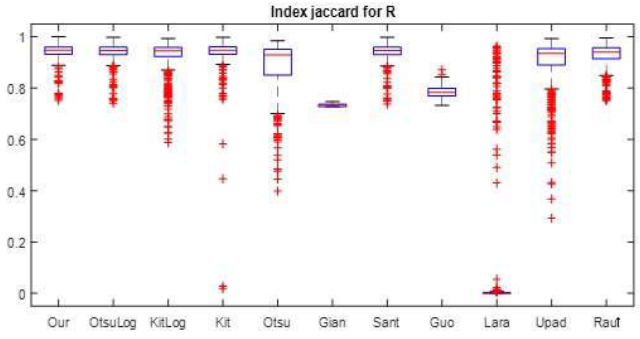

FIGURE 14. Box-plots for the Jaccard index for 11 algorithms tested on red channel of the STARE database images.

TABLE 17. Generalization of experimental results.

Dataset	Kaggle-2015	DDR	RFMID	STARE
Best representation	Red channel	Red channel	Red channel	Red channel
Best algorithms	1.Otsu(Log) 2.Santhakumar **3.Our Alg.** 4.Kittlet(Log) 5.Raut	1.Otsu(Log) 2.Santhakumar **3.Our Alg.** 4.Kittlet(Log) 5.Raut	1.Otsu(Log) 2.Santhakumar **3.Our Alg.** 4.Kittlet(Log) 5.Raut	**1.Our Alg.** 2.Otsu(Log) 3.Santhakumar 4.Kittlet(Log) 5.Raut

V. CONCLUSION

In this paper, we propose a universal method for FOV segmentation based on the ideas of histograms and thresholds. In combination with retinal image representations, we compared 35 segmentation variants.

All of them were tested on 7,000 images selected from the nine largest open-access databases: Kaggle-2015 [4], STARE [18], DIARETDB0 [19], DRIMDB [20], DMED, MESSIDOR, RFMiD, DDR, HRF. The tested images were generated by 100 different fundus cameras in different conditions producing various image size.

Santhakumar R. et al. [16] calculated the binarization threshold of the retinal image on the logarithmic representation of the green channel luminances. Our research proves that these calculations are best done on the red channel. Our algorithm gives very similar results, but it contains 2 empirical parameters, while Otsu's algorithm does not have them and is available in many libraries. Therefore, we recommend to use the following variant as a universal method for segmenting the FOV of areas on color images of the retina formed by an arbitrary fundus camera algorithm No.4 Otsu(Log) from Table 2:

- Perform segmentation of the FOV area on the red channel.
- Convert it into a logarithmic representation.
- Generate the FOV mask by calculating the global threshold by the Otsu method.

This method allows automatic segmentation of the FOV area in the retinal image of different quality, regardless of the type of fundus camera. This is very important when diagnosing diseases based on retinal images using telemedicine.

REFERENCES

[1] K. L. Ong, L. K. Stafford, and S. A. McLaughlin, "Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: A systematic analysis for the Global Burden of Disease Study 2021," *Lancet*, vol. 402, pp. 203–234, Oct. 2023, doi: 10.1016/S0140-6736(23)01301-6.

[2] S. A. Kaveeshwar and J. Cornwall, "The current state of diabetes mellitus in India," *Australas. Med. J.*, vol. 7, no. 1, pp. 45–48, 2014, doi: 10.4066/AMJ.2013.1979.

[3] K. Ogurtsova, L. Guariguata, and N. C. Barengo, "IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021," *Diabetes Res. Clin. Pract.*, vol. 183, Sep. 2022, Art. no. 109118, doi: 10.1016/j.diabres.2021.109118.

[4] F. M. Islam, R. Kawasaki, and R. P. Finger, "Factors associated with participation in a diabetic retinopathy screening program in a rural district in Bangladesh," *Diabetes Res. Clin. Pract.*, vol. 144, pp. 111–117, Oct. 2018.

[5] S. Vujosevic, "A decade-long telemedicine screening program for diabetic retinopathy in the northeast of Italy," *J. Diabetes Complications*, vol. 31, no. 8, pp. 1348–1353, 2017, doi: 10.1016/j.jdiacomp.2017.04.010.

[6] N. Panwar, "Fundus photography in the 21st century—A review of recent technological advances and their implications for worldwide healthcare," *Telemedicine e-Health*, vol. 22, no. 3, pp. 198–208, 2016.

[7] Diabetic Retinopathy Detection. *Kaggle*. Accessed: Mar. 6, 2024. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection

[8] U. Ishtiaq, S. A. Kareem, and E. R. M. F. Abdullah, "Diabetic retinopathy detection through artificial intelligent techniques: A review and open issues," *Multimedia Tools Appl.*, vol. 79, no. 21, pp. 15209–15252, 2020, doi: 10.1007/s11042-018-7044-8.

[9] V. V. Starovoitov, I. Y. Golub, and M. M. Lukashevich, "A universal retinal image template for automated screening of diabetic retinopathy," *Pattern Recognit. Image Anal.*, vol. 32, no. 2, pp. 322–331, 2022, doi: 10.1134/s1054661822020195.

[10] G. Kovács and A. Fazekas, "A new baseline for retinal vessel segmentation: Numerical identification and correction of methodological inconsistencies affecting 100+ papers," *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102300, doi: 10.1016/j.media.2021.102300.

[11] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Sep. 1979, doi: 10.1109/TSMC.1979.4310076.

[12] B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imag.*, vol. 13, no. 1, p. 146, 2004, doi: 10.1117/1.1631315.

[13] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986, doi: 10.1016/0031-3203(86)90030-0.

[14] V. V. Starovoitov, Y. I. Golub, and M. M. Lukashevich, "Digital fundus image quality assessment," *Syst. Anal. Appl. Inf. Sci.*, vol. 4, pp. 25–38, Jan. 2022, doi: 10.21122/2309-4923-2021-4-25-38.

[15] LgiancaUTH/hei-med. *The Hamilton Eye Institute Macular Edema Dataset (Hei-Med)*. Accessed: Mar. 6, 2024. [Online]. Available: https://github.com/lgiancaUTH/HEI-MED
[16] R. Santhakumar, E. R. Rajkumar, M. Tandur, K. T. Rajamani, and G. Haritz, "Novel method for automatic generation of fundus mask," in *Proc. 3rd Int. Conf. Image Inf. Process. (ICIIP)*, Dec. 2015, pp. 147–151, doi: 10.1109/ICIIP.2015.7414756.
[17] J. Guo, G. Azzopardi, C. Shi, N. M. Jansonius, and N. Petkov, "Automatic determination of vertical Cup-to-Disc ratio in retinal fundus images for glaucoma screening," *IEEE Access*, vol. 7, pp. 8527–8541, 2019, doi: 10.1109/ACCESS.2018.2890544.
[18] L. D. L. Rodriguez and G. U. Serrano, "Exudates and blood vessel segmentation in eye fundus images using the Fourier and cosine discrete transforms," *Computacion y Sistemas*, vol. 20, no. 4, pp. 697–708, 2016, doi: 10.13053/cys-20-4-2305.
[19] K. Upadhyay, M. Agrawal, and P. Vashist, "Unsupervised multiscale retinal blood vessel segmentation using fundus images," *IET Image Process.*, vol. 14, no. 11, pp. 2616–2625, 2020, doi: 10.1049/iet-ipr.2019.0969.
[20] R. Raut, "Laser scar classification in retinal fundus images using wavelet transform and local variance," in *Computer Vision and Machine Intelligence in Medical Image Analysis*, vol. 992. Singapore: Springer, 2020, pp. 81–90.

V. V. STAROVOITOV received the Doctor of Sciences degree. He is currently a Professor in computer science. He is also a Principal Research Fellow with the United Institute of Informatics Problems, National Academy of Sciences of Belarus (UIIP NAN Belarus). He has published over 150 articles. His research interests include Prof. Starovoitov are processing and analysis of digital images obtained in different parts of the electromagnetic spectrum. He received the Award: The State Prize of the Republic of Belarus in Science.

NGUYEN NHU SON received the Ph.D. degree in computer science from The University of Queensland, Australia, in 2007. He is currently the Vice Chair of Scientific Council and the Head of the Department with the Institute of Information Technology (IOIT), Vietnam Academy of Science and Technology. His research interests include artificial intelligence, data mining, soft computing, and fuzzy computing.

YU. I. GOLUB received the Ph.D. degree. She is currently an Associate Professor. She is also a Senior Research Fellow with the United Institute of Informatics Problems, National Academy of Sciences of Belarus. Her main research interests include digital image processing and image quality assessment.

M. M. LUKASHEVICH received the Ph.D. degree. She is currently an Associate Professor. She is also a Postdoctoral Researcher with Belarusian State University of Informatics and Radioelectronics. Her main research interests include machine learning and computer vision.

NGUYEN LONG GIANG received the Ph.D. degree in mathematics from the Institute of Information Technology (IoIT), Vietnam Academy of Science and Technology, in 2012. He is currently an Associate Professor and the Vice Director of IoIT, Vietnam Academy of Science and Technology. His research interests include artificial intelligence, data mining, soft computing, and fuzzy computing. He serves as a Chief of Editor for the *Vietnam Journal of Computer Science and Cybernetics* (JCC).

HOANG THI MINH CHAU received the master's degree from the Faculty University Natural Studies, VNU, in 2010. She is currently pursuing the Ph.D. degree with the Institute of Information Technology, Vietnam Academy of Sciences. She is also a Lecturer with the Information Technology Department, University of Economics-Technology for Industries. Her research interests include the IoT, AI, data mining data, and machine learning.

LE HOANG SON received the joint Ph.D. degree in mathematics and informatics from the VNU University of Science, Vietnam National University (VNU), in conjunction with Politecnico di Milano, Italy, in 2013. He has been promoted to an Associate Professor in information technology in Vietnam, since 2017. He is currently the Vice Director of the VNU Information Technology Institute. His major research interests include artificial intelligence, data mining, soft computing, fuzzy computing, fuzzy recommender systems, and geographic information systems. He is an Associate Editor of the *Journal of Intelligent and Fuzzy Systems* (SCIE), IEEE Access (SCIE), *Data Technologies and Applications* (SCIE), *International Journal of Data Warehousing and Mining* (SCIE), *Neutrosophic Sets and Systems* (ESCI), *Vietnam Research and Development on Information and Communication Technology*, *VNU Journal of Science: Computer Science and Communication Engineering*, and *Frontiers in Artificial Intelligence*.