

# A Decision Support System for Early-Stage Diabetic Retinopathy Lesions

Kemal AKYOL

Computer Engineering Department  
Kastamonu University  
Kastamonu, Turkey

Şafak BAYIR

Computer Engineering Department  
Karabuk University  
Karabuk, Turkey

Baha ŞEN

Computer Engineering Department  
Yildirim Beyazıt University  
Ankara, Turkey

**Abstract**—Retina is a network layer containing light-sensitive cells. Diseases that occur in this layer, which performs the eye-sight, threaten our eye-sight directly. Diabetic Retinopathy is one of the main complications of diabetes mellitus and it is the most significant factor contributing to blindness in the later stages of the disease. Therefore, early diagnosis is of great importance to prevent the progress of this disease. For this purpose, in this study, an application based on image processing techniques and machine learning, which provides decision support to specialist, was developed for the detection of hard exudates, cotton spots, hemorrhage and microaneurysm lesions which appear in the early stages of the disease. The meaningful information was extracted from a set of samples obtained from the DIARETDB1 dataset during the system modeling process. In this process, Gabor and Discrete Fourier Transform attributes were utilized and dimension reduction was performed by using Spectral Regression Discriminant Analysis algorithm. Then, Random Forest and Logistic Regression and classifier algorithms' performances were evaluated on each attribute dataset. Experimental results were obtained using the retinal fundus images provided from both DIARETDB1 dataset and the department of Ophthalmology, Ataturk Training and Research Hospital in Ankara.

**Keywords**—Early stage diabetic retinopathy lesions; feature extraction; important features; image recognition; classification; decision support system; computer aided analysis

## I. INTRODUCTION

Diabetic Retinopathy (DR) which is the subject of many studies in medical image processing field is a disease that begins with the influence of the retinal capillaries due to effect of blood sugar increase depending on diabetes and can result complete loss of sense of sight in its progressive stages [1]-[2]. There are two phases of DR disease which is directly proportional with the level of structural deterioration in retinal images: early-stage diabetic retinopathy (ESDR) and advanced-stage diabetic retinopathy (ASDR) [3]. Clogging of network layer vessels, small vessel dilatations, intraretinal hemorrhages and yellow deposits called hard exudate are seen at the onset of this disease [2], [4]. Lesion samples which occur in the early stage of this disease are given in Fig. 1.

Automatic detection and segmentation studies on ESDR disease gained great momentum in recent years, furthermore, new competencies are being added each passing day. In addition, the regular examination requirement of this disease

and the lack of specialist make the procedures that should be carried out with automated systems compulsory.

The aim of this study is to investigate the methodology and techniques that will enable us to detect accurately the location of the structural disorders namely lesions occur in the early stage of DR and to model the decision support system that gives the most accurate result.

The application of ESDR lesions' detection which is based on literature reviews and reference to the tissue classification approach, involves the construction of the model and the analysis of new retinal images basically. This application includes the interfaces that are useful for field specialists in the decision of improving their cognitive abilities related to understanding and comprehension.

**Contribution:** There are many studies on the detection of ESDR lesions in the relevant literature. This study shows similarities with other studies in terms of workflow, but it differs from others using Discrete Fourier Transform Attributes (DFTA) and Spectral Regression Discriminant Analysis (SRDA) algorithms. On the other hand, the saving ability for the ESDR lesions position in retinal images which are taken at different dates of a patient, allows the field specialist to instantly compare the patient's previous recordings in the system, thereby allows specialist in order to examine the development of disease in particular date and time intervals.

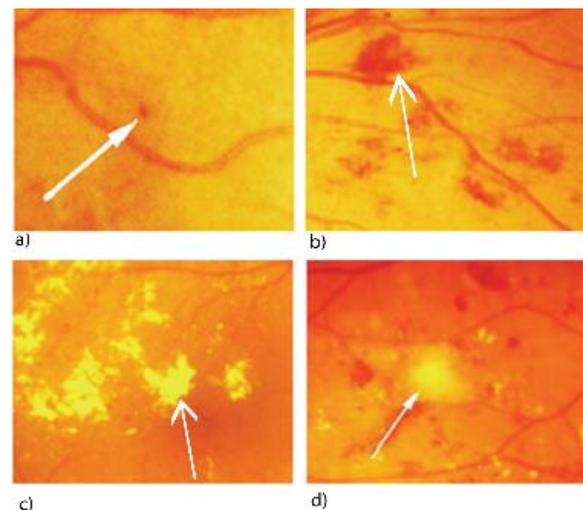


Fig. 1. ESDR structural disorders, a) microaneurysm, b) hemorrhage, c) hard exudate, d) cotton spots [5].

The rest of this paper is organized as follows: related literature and studies were presented in Section 2. In Section 3, the proposed method was handled in a detailed way. In Section 4, the developed application and experimental results were presented. Finally, conclusions and future studies were given in Section 5.

## II. LITERATURE REVIEW

Some of the studies in the literature on this subject that may result in complete loss of sense of sight in the progressive stages, are as follows: Kumar et al. classified the attributes which were extracted from the obtained regions by using morphological pre-processing, image boundary monitoring and Otsu thresholding techniques, with Support Vector Machines (SVM) in their studies [6]. Quellec et al. trained the system for recognition of patterns having random dimensions marked by doctors for automatic image classification, and then precisely identified the DR lesions by classifying the similar pattern images as lesions or not [7]. Mookiah et al. presented automatic scanning system for early stage and advanced stage diabetic retinopathy, and normal images. They proposed a system involving the processing of eyebrow images for the extraction of abnormal signals. They worked with probabilistic neural networks, decision tree and SVM, and utilized from statistically significant 13 features for classification. Also, they used Particle Swarm Optimization and Genetic Algorithm Optimization techniques in their studies [8]. Askew et al. presented a study, aiming to assess the detection of DR's early stage, management and its benefits, handling situations such as patient information, general practitioners, ophthalmologists, screening rates and monitoring of the early stage of disease appropriately [9]. Niemeijer et al. used the machine learning technique which they called it as the supervised algorithm in order to detect bright lesions and determine how to distinguish between lesions [10]. Garcia et al. performed automatic detection of hard exudate regions by using the feature set that distinguish hard exudate regions best in retinal images from each other [11]. Hipwell et al. categorized the microaneurysm candidate regions using intensity and dimension information in the framework of the rules that they obtained as a result of training 102 images [12]. Sopharak et al. performed the detection of microaneurysms by using morphology, segmentation and naive Bayes classifier [13]. Sharma et al. implemented a dynamic thresholding algorithm based on image processing techniques for the detection of hemorrhage regions in retinal images. They determined these regions by using the color and size information of hemorrhage regions [14]. Saleem and Usman Akram proposed the detection of hemorrhage lesions based on color features including pre-processing, light thresholding, extraction of candidate regions, attribute extraction and classification stages [15]. Spencer et al. performed thresholding on the image they obtained by applying a bilinear top-hat morphological transformation and then matched filtering, and obtained binary images. Afterwards, they carried out the regional enlargement process by examining the regions for the detection of microaneurysms lesions [16]. As can be seen from these studies, the lesioned regions were detected by using various distinguishing features and different classifier algorithms on the images processed by

applying the basic image processing algorithms and morphological operations.

## III. DATA AND METHODOLOGY

### A. Data

Retinal images used in this study were taken from Ankara Ataturk Training and Research Hospital Ophthalmology Department and publicly available DIARETDB1 (Standard Diabetic Retinopathy Database) datasets were used. The images in first dataset are 2304x1536 and the others are 1500x1152 in size, and they have all 24-bit depth.

### B. Methodology

1) *Image Enhancement*: After the digital images are obtained, they are pre-processed. This step is a crucial step that affects the performance of computer vision systems, and the critical decision making about the image. It is aimed to obtain the best performance from processed image by the image enhancement which is one of the pre-processing stage.

2) *Detection of Interest Points, and Attribute Extraction*: The interest points are the points that give the greatest reaction in the regions where the change occurs in an image. As can be seen in the literature studies [17]-[20], the meaningful information is obtained from the region of interest around these points. The answer to the question "Which one is the most accurate algorithm?" is searched at the stage when constructing the attributes vector that represents the region of interest best, and this stage is also called as the attribute extraction. As a result of long-term experimental studies, the interest points are extracted with Oriented FAST and Rotated BRIEF, which are an Oriented and Rotated Binary Robust Independent Basic Feature algorithms [21], and SURF (Speeded Up Robust Features) [22] which is an accelerated robust feature algorithm widely used in the literature. In order to extract the meaningful information from these points of interest, the Gabor Attributes (GA) and DFTA were used.

3) *Dimension Reduction*: The dimension reduction method which aims to reduce a dataset to a smaller size that represents its original set, is a well-known subject in machine learning and there are numerous studies on this subject. Decreasing data size through reducing the attributes will provide faster execution times of the classification algorithms which perform numerous identification process. But, ideally, it should protect the essence information which have high discrimination power and high reliability in order to obtain the better performance [23]. Various dimension reduction algorithms were tried and finally Spectral Regression Discriminant Analysis (SRDA) algorithm was decided to be used in experimental studies. SRDA [24], an efficient and a new method that obtains transformation vectors from a set of linear regression problems, includes graph-based formulations of linear discriminant analysis that is one of the dimension reduction methods [24], [25].

4) *Classification and Learning*: In this process which is a necessary and an important stage for machine learning, it is aimed to analyse the region of interest accurately. For this reason, the learning process including the training and testing stages is performed. The attribute dataset and actual result values of the sample set are given as input, and the system learns which output should be generated for this dataset during the training phase. Thus, the system decides the class of the analyzed data in the light of previous information. All data included in the classification phase is passed through the same workflow, and so the performance of each classifier is analyzed. Several classifier algorithms were used, and finally LR and RF algorithms were continued to be used for application developed throughout the study.

5) *Performance Evaluation*: K-fold cross-validation technique is commonly used in order to make test results much more reliable and determinative in many classification studies. Training and test data are crossly replaced, therefore, the mistakes associated with the random sampling of training set are minimized and machine memorization is blocked with this technique. That is, all data is divided into approximately equal sub-datasets determined with the k, and then the classifier algorithm is trained and tested k times. While taking one of the folds each time as a test data, the others are formed for training data [26]. Accuracy analysis is a process which is aimed to determine the correctness of the classes assigned by the classifier algorithms. The success of the algorithms is evaluated by comparing the observed results and predicted results [27]. Regarding the processing region of interest, “positive” observation refers to structural disorder (1), “negative” observation refers to structural disorder is not (0), and the situations that can be encountered in this process are as follows:

- TP is the number of lesioned region found as positive,
- TN is the number of normal region found as negative,
- FP is the number of normal region found as positive,
- FN is the number of lesioned region found as negative.

The standard performance measures, Sensitivity (Sen), Specificity (Spe) and Accuracy (Acc) which was given in (1), (2) and (3), respectively were used in order to evaluate the performance of the learning algorithms.

$$\text{Sen} = \text{TP} / (\text{TP} + \text{FN}) \quad (1)$$

$$\text{Spe} = \text{TN} / (\text{TN} + \text{FP}) \quad (2)$$

$$\text{Acc} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (3)$$

The Sen is the ratio of the number of actual correct positives in total positives. That is, the percentage of positives were classified as positive. The Spe is the ratio of the number of actual correct negatives in total negatives. The Acc value which is the most commonly used in measurement of model success, can be expressed as the ratio of the number of accurately diagnosed samples to the number of total samples

[27]. Improvements are made considering in each process in the workflow if these values are not in ideal level.

## IV. DEVELOPED APPLICATION AND EXPERIMENTAL RESULTS

### A. Model Building Process

The developed application and experimental studies were carried out in the Python 2.7 platform in this study. Software codes were developed including “scikit-image” library for image processing, and “scikit-learn” and “mlpy” machine learning libraries for learning. In order to store the data and use these stored data in the related interfaces, the databases and tables were prepared within entity-relationship model framework with MySQL software. The general workflow diagram of the developed application is given in Fig. 2.

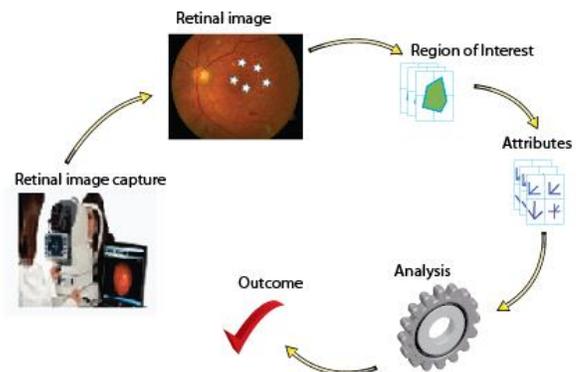


Fig. 2. The working scheme of the developed application.

There are some noisy and uneven illuminations on the raw-retinal images because of the light which is falling on the retina layer at different density. For this reason, as seen in the application steps in Fig. 3, it is aimed to improve the image quality by applying “whitening” and CLAHE pre-processing techniques on retinal image in the RGB color space, and thus making the image ready for analysis.

The learning process of the developed system was designed according to the workflow in Fig. 4. In this process, the answer to the question: “Which is the most successful model?” was sought. The color intensity of hemorrhage and microaneurysm lesions in the retinal image is approximately the same as seen in Fig. 3(a) and 3(b). The color intensity of hard exudate and soft exudate lesions are approximately same. For this reason, two group sample sets were prepared by 150 regions of interest information, including 75 positive and negative interest regions in random dimensions, were grouped as the first group lesions for hard and soft exudates, and as the second group lesions for hemorrhage and microaneurysm. The GA and DFTA datasets obtained from both sample image sets were [150x512] and [150x400] dimensions, respectively. Dimension reductions were made on these datasets with the SRDA algorithm. 80% and 20% of these datasets were reserved for training and testing respectively within the context of 5-fold cross-validation technique. In this respect, the results belong to each built model, which were represented in confusion matrix format, were presented from Tables I to IV. The averages of the experimental results obtained were given at the bottom of the related tables.

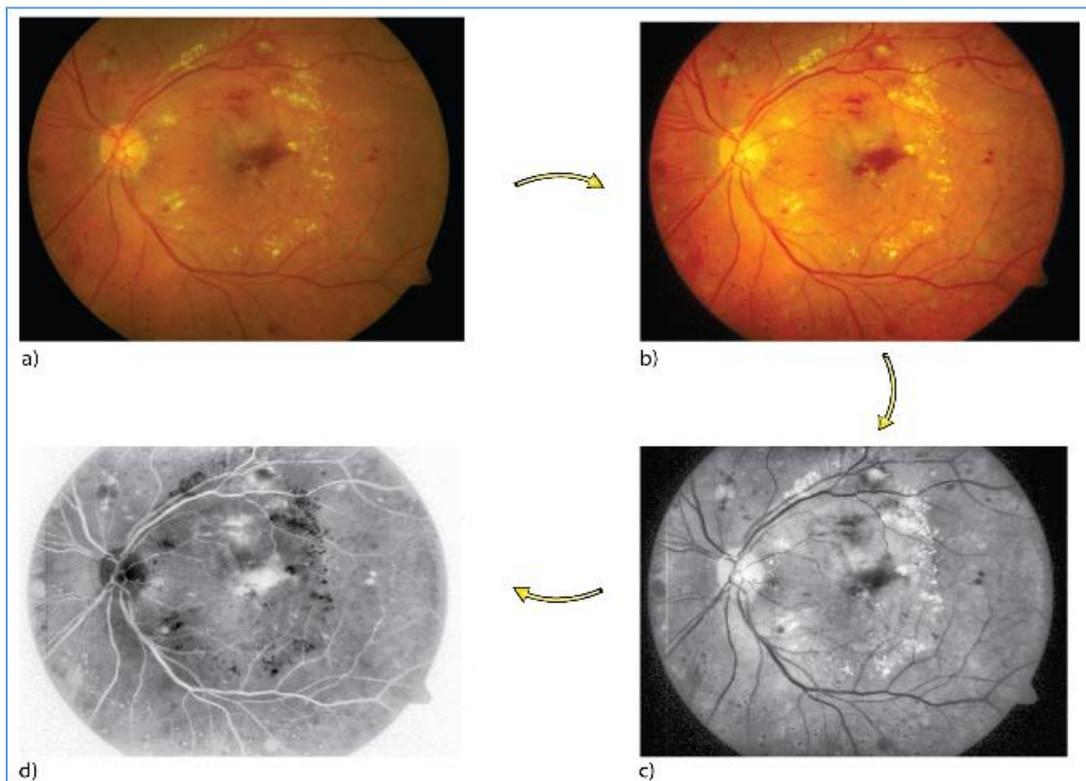


Fig. 3. Pre-processing process: a) RGB image; b) Improvement of RGB image by whitening method (pre-processing -1); c) Improvement of Fig 3(b) image by CLAHE method (pre-processing -2); d) the reverse of Fig. 3(c) image.

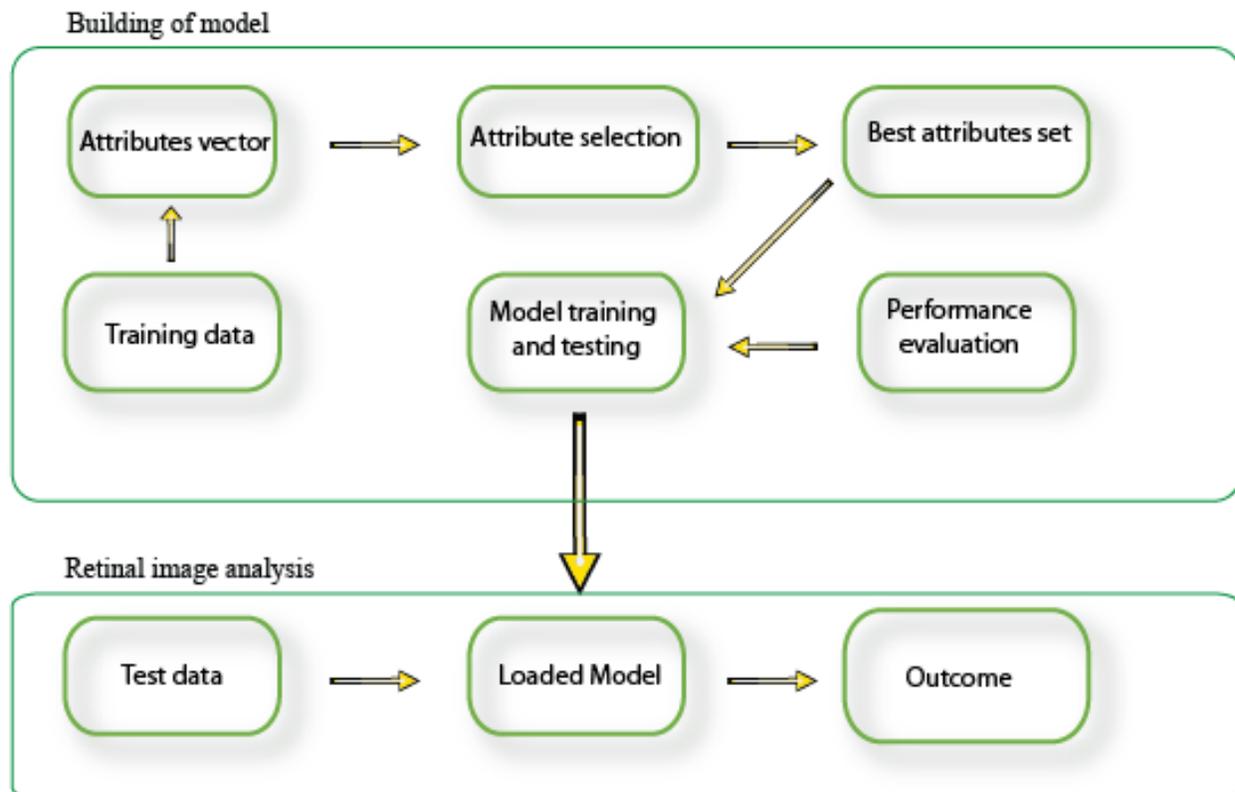


Fig. 4. The workflow of the learning process.

TABLE I. CLASSIFICATION RESULTS FOR HARD EXUDATE AND COTTON SPOTS WITH DFTA AND SRDA

<b>DFTA + SRDA</b>						
<i>Classifier Algorithms</i>						
<b>Datasets</b>	<b>LR</b>			<b>RF</b>		
Dataset No:1		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	10	5	<b>1</b>	11	4
	<b>0</b>	0	15	<b>0</b>	0	15
		Acc	<b>%83.33</b>		Acc	<b>86.66</b>
Dataset No:2		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	12	3	<b>1</b>	13	2
	<b>0</b>	0	15	<b>0</b>	1	14
		Acc	<b>%90.0</b>		Acc	<b>90.0</b>
Dataset No:3		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	14	1
	<b>0</b>	1	14	<b>0</b>	0	15
		Acc	<b>93.33</b>		Acc	<b>96.66</b>
Dataset No:4		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	13	2	<b>1</b>	13	2
	<b>0</b>	1	14	<b>0</b>	2	13
		Acc	<b>90.0</b>		Acc	<b>86.66</b>
Dataset No:5		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	13	2
	<b>0</b>	1	14	<b>0</b>	1	14
		Acc	<b>93.33</b>		Acc	<b>90.0</b>
Average Acc	<b>% 90.0</b>			<b>% 90.0</b>		

TABLE II. CLASSIFICATION RESULTS FOR HARD EXUDATE AND COTTON SPOTS WITH GA AND SRDA

<b>GA and SRDA</b>						
<i>Classifier Algorithms</i>						
<b>Datasets</b>	<b>LR</b>			<b>RF</b>		
Dataset No:1		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	9	6	<b>1</b>	12	3
	<b>0</b>	7	8	<b>0</b>	3	12
		Acc	<b>56.66</b>		Acc	<b>80.0</b>
Dataset No:2		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	10	5
	<b>0</b>	3	12	<b>0</b>	3	12
		Acc	<b>86.66</b>		Acc	<b>73.33</b>
Dataset No:3		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	9	6	<b>1</b>	10	5
	<b>0</b>	1	14	<b>0</b>	0	15
		Acc	<b>76.66</b>		Acc	<b>83.33</b>
Dataset No:4		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	11	4	<b>1</b>	10	5
	<b>0</b>	1	14	<b>0</b>	3	12
		Acc	<b>83.33</b>		Acc	<b>73.33</b>
Dataset No:5		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	10	5	<b>1</b>	10	5
	<b>0</b>	4	11	<b>0</b>	3	12
		Acc	<b>70.0</b>		Acc	<b>73.33</b>
Average Acc	<b>% 74.66</b>			<b>% 76.66</b>		

TABLE III. CLASSIFICATION RESULTS FOR HAEMORRHAGE AND MICROANEURYSM WITH DFTA AND SRDA

DFTA and SRDA						
Classifier Algorithms						
Datasets	LR			RF		
Dataset No:1		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	15	0	<b>1</b>	15	0
	<b>0</b>	3	12	<b>0</b>	1	14
		Acc	<b>90.0</b>		Acc	<b>%96.66</b>
Dataset No:2		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	15	0	<b>1</b>	14	1
	<b>0</b>	3	12	<b>0</b>	2	13
		Acc	<b>90.0</b>		Acc	<b>% 90.0</b>
Dataset No:3		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	13	2
	<b>0</b>	2	13	<b>0</b>	2	13
		Acc	<b>% 90.0</b>		Acc	<b>%86.66</b>
Dataset No:4		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	13	2
	<b>0</b>	3	12	<b>0</b>	1	14
		Acc	<b>%86.66</b>		Acc	<b>% 90</b>
Dataset No:5		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	11	4
	<b>0</b>	3	12	<b>0</b>	2	13
		Acc	<b>%86.66</b>		Acc	<b>%80.0</b>
Average Acc	<b>88.66</b>			<b>88.66</b>		

TABLE IV. CLASSIFICATION RESULTS FOR HAEMORRHAGE AND MICROANEURYSM WITH GA AND SRDA

GA and SRDA						
Classifier Algorithms						
Datasets	LR			RF		
Dataset No:1		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	10	5
	<b>0</b>	3	12	<b>0</b>	5	10
		Acc	<b>86.66</b>		Acc	<b>66.66</b>
Dataset No:2		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	12	3
	<b>0</b>	2	13	<b>0</b>	0	15
		Acc	<b>90.0</b>		Acc	<b>90.0</b>
Dataset No:3		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	13	2
	<b>0</b>	4	11	<b>0</b>	4	11
		Acc	<b>83.33</b>		Acc	<b>80.0</b>
Dataset No:4		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	12	3	<b>1</b>	11	4
	<b>0</b>	5	10	<b>0</b>	5	10
		Acc	<b>73.33</b>		Acc	<b>70.0</b>
Dataset No:5		<b>1</b>	<b>0</b>		<b>1</b>	<b>0</b>
	<b>1</b>	14	1	<b>1</b>	13	2
	<b>0</b>	3	12	<b>0</b>	3	12
		Acc	<b>86.66</b>		Acc	<b>83.33</b>
Average Acc	<b>% 84.0</b>			<b>% 78.0</b>		

In addition, Sn, Sp and Acc values were presented from Tables V to VIII. According to these results, it can be said that

the hybrid combination of DFTA feature extraction, SRDA feature reduction methods and, RF and LR classifier algorithms are very successful.

TABLE V. EVALUATION OF MODEL 1 PERFORMANCE FOR 1ST GROUP

Datasets	Model 1 (DFTA + SRDA + LR)		
	Sen %	Spe %	Acc %
Dataset No:1	66.66	100.0	83.33
Dataset No:2	80.0	100.0	90.0
Dataset No:3	93.33	93.33	93.33
Dataset No:4	86.66	93.33	90.0
Dataset No:5	93.33	93.33	93.33
Average	84.0	96.0	90.0

TABLE VI. EVALUATION OF MODEL 2 PERFORMANCE FOR 1ST GROUP

Datasets	Model 2 (DFTA + SRDA + RF)		
	Sen %	Spe %	Acc %
Dataset No:1	73.33	100.0	86.66
Dataset No:2	86.66	93.33	90.0
Dataset No:3	93.33	100.0	96.66
Dataset No:4	86.66	86.66	86.66
Dataset No:5	86.66	93.33	90.0
Average	85.33	94.66	90.0

TABLE VII. EVALUATION OF MODEL 1 PERFORMANCE FOR 2ND GROUP

Datasets	Model 1 (DFTA + SRDA + LR)		
	Sen %	Spe %	Acc %
Dataset No:1	100.0	80.0	90.0
Dataset No:2	100.0	80.0	90.0
Dataset No:3	93.33	86.66	90.0
Dataset No:4	93.33	80.0	86.66
Dataset No:5	93.33	80.0	86.66
Average	96.0	81.33	88.66

TABLE VIII. EVALUATION OF MODEL 2 PERFORMANCE FOR 2ND GROUP

Datasets	Model 2 (DFTA + SRDA + RF)		
	Sen %	Spe %	Acc %
Dataset No:1	100.0	93.33	96.66
Dataset No:2	93.33	86.66	90.0
Dataset No:3	86.66	86.66	86.66
Dataset No:4	86.66	93.33	90.0
Dataset No:5	73.33	86.66	80.0
Average	88.0	89.33	88.66

### B. Analysis Process of Retinal Images

As seen in the flowchart in Fig. 5, the analysis process of a retinal fundus image is as follows: Firstly, bright regions extraction algorithm was applied to both image; the first image was obtained by “whitening” and CLAHE operations, and the second image was obtained by reversing this image, in order to obtain the 1st and 2nd group lesioned regions. After this step, the set of points of interest obtained by both ORB and SURF algorithms were determined. Automated detection of lesioned regions in retinal image was performed after manual detection of the optic disc and macula regions, and blood vessels detection. Since the number of interest points directly affect the execution time of the analysis, Analysis of the points of interest was carried out with an equally shared 4-threaded structure. The hard or soft exudate structural disorder is assigned for the 1st group lesion type, and a hemorrhage or microaneurysm

structural disorder is assigned for the 2nd group lesion type. Hard and soft exudate lesions are included in the same category due to color intensity and brightness ratio similarity. On the other hand, the hemorrhage and microaneurysm regions were assigned to 2nd group which were distinguished with the model based on the DFTA knowledge. Image recording is used to help the doctors who make treatment decisions in monitoring the development of any disease in medical field. In this context, a recording and tracking modules were developed within the settings of the steps mentioned above. The abilities of these modules are as follows:

- The lesions’ analysis.
- Inputting of the related data for the image analysed.
- Monitoring the course of the disease.

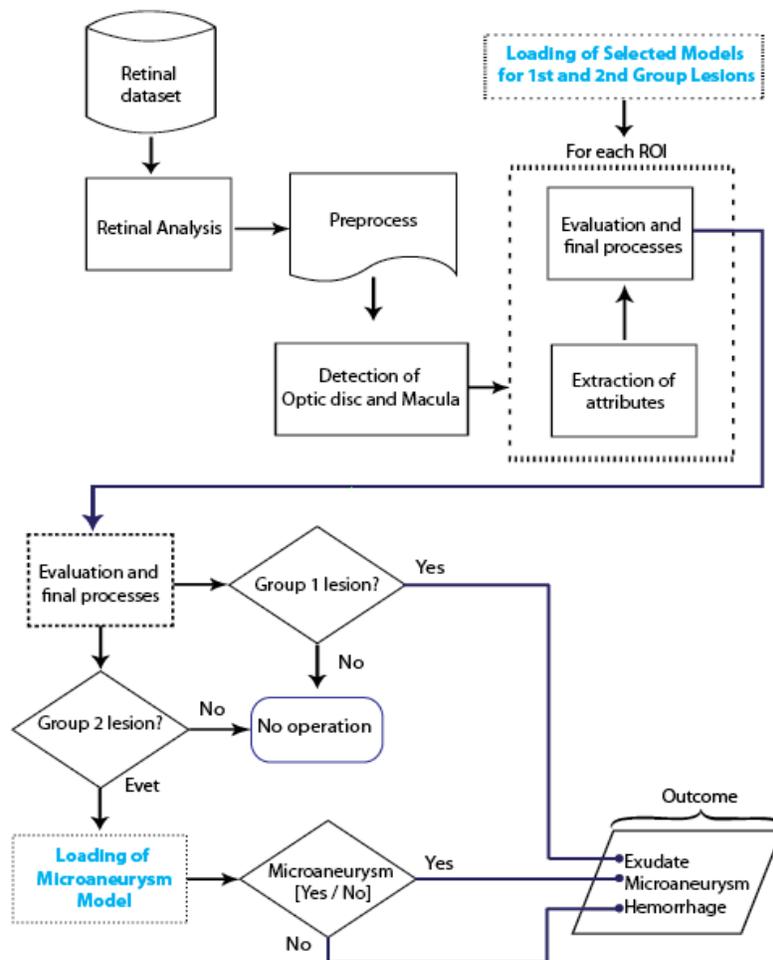


Fig. 5. General block diagram of the process of analysis of new retinal images.

The result image obtained with the algorithm which performs the detection of the 1st group exudate lesions is seen in Fig. 6. The regions in blue color denoted lesions in this figure. A field specialist could make the following steps on these images analysed.

- Examining the regions with lesion or other regions (Fig. 6).

- Removing the particular interest region (Fig. 6).
- Defining the 'x' kind of lesion have or not in the region (Fig. 6).
- Examining the image in more detail through the interface (Fig. 7).

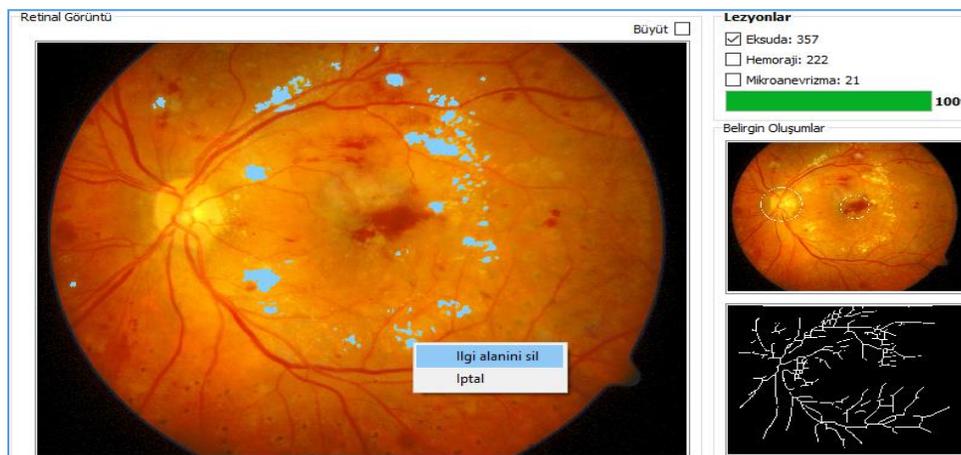


Fig. 6. Update of lesioned regions (Update of the regions of interest).

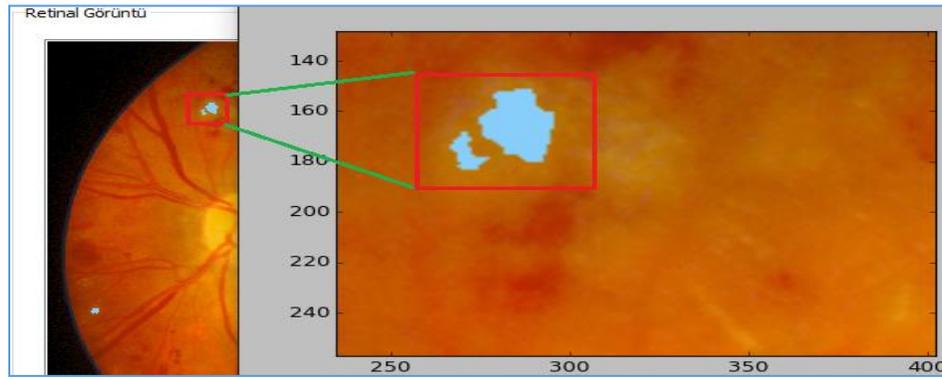


Fig. 7. Detailed examination of the lesioned region.

The exemplary output of the vein subtraction algorithm which was carried out for not analyzing the points of interest on the blood vessels is seen in Fig. 8. The pseudo code that expresses instructions mentioned above was given below. All the information obtained from the analysis are recorded to the “patient information” database designed within the context of the entity-relationship model and after, these records could be easily accessed. On the other hand, ESDR suspected persons are notified by the system in certain intervals.

```
def is_interest_point_over(point_x,point_y):  
    load_vein_structure(retina_id)  
    is_interest_point_on_vein_structure?  
    If the answer is yes, remove this point from the analys_list.
```

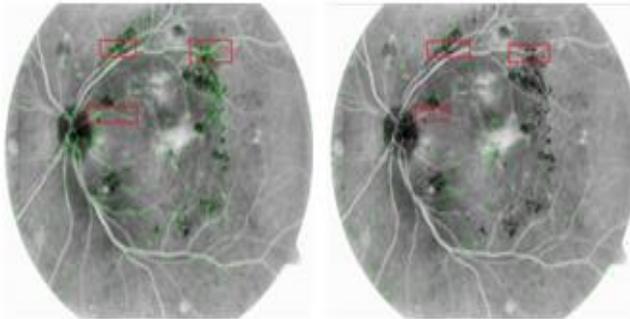


Fig. 8. The detection of points of interest on the blood vessels. a) SURF keypoints b) ORB keypoints.



Fig. 9. The factors that increase the DR risk to be recorded in the system.

Also, factors listed below that increase the risk of DR were included in the developed system in the direction of [3].

- The period of diabetes illness.
- High blood sugar level.
- High blood pressure.
- Pregnancy.
- Obesity.
- Kidney disease.

The field specialist examines the risk factor values on the selected date and time, and makes inference by utilizing the information of red bordered region in Fig. 9. The most important of enrollment comparison screen (Fig. 10) task is to provide the necessary information about the disease progress, which are listed below, to the field specialist:

- Accessing the records of the relevant retinal images existed within the system.
- Performing a retinal image search.
- Reviewing lesioned regions.
- Reviewing and comparing the retinal images processed in two different dates or times.

For example, the lesions in red colour Fig. 10 show new findings and the green ones show previously recorded findings.

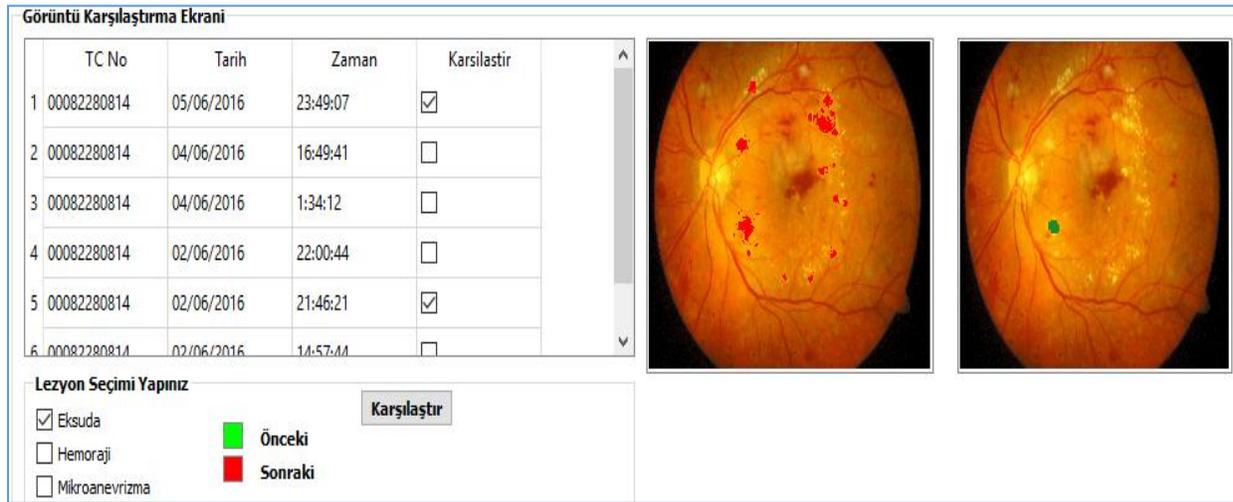


Fig. 10. Examination of two registered data in the system.

## V. CONCLUSION

DR is a very common eye disease that starts with being effected of the retinal capillary vessels as a result of increase of the blood sugar, and can result in complete loss of sense of sight in its progressive stages and usually affects both eyes in the blindness degree. The most important method for preventing this disease is early diagnosis. Within this context, an application based on machine learning and image processing techniques, which provides a decision support to field specialist for the detection of hard exudate, soft exudate, hemorrhage and microaneurysm lesions that emerge in the early stages of this disease was developed. The learning process of the developed application consists of these stages: enhancement of retinal images, the detection of region of interests and attributes extraction, and the analysis of these regions. It was seen that the developed application which is the hybrid decision-making system presented successful results for each lesion group in the learning process. But, it is not possible to achieve this success on images with very different color intensity and brightness ratio such as retinal glaucoma, macula diseases or ASDR. On the other hand, the interfaces were designed to enable the field specialist to examine and instantly compare the developments occur in specific dates and times. It is thought that this study provides assistance to the field specialist for the detection of the lesions occurring in ESDR disease, and prepares a ground for the increase in the success of treatment of the patient, and will shed light for future works. In the future, it is aimed to increase the effectiveness of the application by utilizing new attributes and classification algorithms. Moreover, the detection of lesions occurring in the ASDR is also among the targeted studies.

## ACKNOWLEDGMENT

The authors would like to thank to Asst. Prof. Dr. Hilal Kaya and Asst. Prof. Dr. Mücella Arıkan, staff of Ophthalmology Department in Ankara Atatürk Training and Research Hospital, who are helpful for providing retinal images used in the study and also the developers of the DIARETDB1 project who provide the retinal fundus dataset which they created to the researchers' use in their studies.

## REFERENCES

- [1] Internet: NHS Choices your health, your choices, "Diabetic retinopathy", <http://www.nhs.uk/Conditions/Diabetssic-retinopathy/Pages/Introduction.aspx> (2017).
- [2] Internet: İstanbul Cerrahi, "Diyabetik Retinopati", <http://www.lasikinturkey.com/diyabetik-retinopati> (2017).
- [3] Internet: Türk Diabet ve Obezite Vakfı, "Diabetik Retinopati", <http://diabetvakfi.org/inf.php?partid=5&catid=5&pid=26> (2017).
- [4] Q. Mohamed, M. C. Gillies and T. Y. Wong. "Management of diabetic retinopathy: A systematic review," The Journal of the American Medical Association, vol. 298, pp. 902-916, 2007.
- [5] T. Kauppi, J. K. Kamarainen, L. Lensu, V. Kalesnykiene, I. Sorri, H. Uusitalo and H. Kalviainen, "Constructing benchmark databases and protocols for medical image analysis: Diabetic Retinopathy," Computational and Mathematical Methods in Medicine vol. 2013, pp. 1-15, 2013.
- [6] A. Kumar, A. K. Gaur and M. Srivastava, "A segment based technique for detecting exudate from retinal fundus image," Procedia Technology, vol. 6, pp. 1-9, 2012.
- [7] G. Quellec, M. Lamard, M. D. Abramoff, E. Decencière, B. Lay, A. Erginay, B. Cochener and G. Cazuguel, "A multiple-instance learning framework for diabetic retinopathy screening," Medical Image Analysis vol. 16, 1228-1240, 2012.
- [8] M.R.K. Mookiah, U. Rajendra-Acharya, R.J. Martis, K.C. Chua, C.M. Lim, E.Y.K. Ng and A. Laude, "Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach," Knowledge-Based Systems vol. 39, pp. 9-22, 2013.
- [9] D.A. Askew, L. Crossland, R.S. Ware, P.C. StephenBegg, P. Mitchell and L.J. Claire, "Diabetic retinopathy screening and monitoring of early stage disease in general practice: design and methods," Contemporary Clinical Trials vol. 33, pp. 969-975, 2012.
- [10] M. Niemeijer, B. van Ginneken and S.R. Russell, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," Investigative Ophthalmology & Visual Science, vol. 48, pp. 2260-2267, 2007.
- [11] M. Garcia, C. Valverde, M.I. Lopez, J. Poza and R. Hornero, "Comparison of logistic regression and neural network classifiers in the detection of hard exudates in retinal images," 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 3-7 July 2013; Osaka, Japan: pp. 5891-5894.
- [12] J.H. Hipwell, F. Strachan, J.A. Olson, K.C. McHardy, P.F. Sharp and J.V. Forrester, "Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool," Diabet Med, vol. 17, pp. 588-594, 2000.

- [13] A. Sopharak, B. Uyyanonvara, S. Barman, "Simple hybrid method for fine microaneurysm detection from non-dilated diabetic retinopathy retinal images," *Computerized Medical Imaging and Graphics*, vol. 37, pp. 394-402, 2013.
- [14] A. Sharma, M.K. Dutta, A. Singh, M. Parthasarathi and C.M. Travieso, "Dynamic thresholding technique for detection of hemorrhages in retinal images," *Seventh International Conference on Contemporary Computing*; 7-9 Aug 2014; Noida, India.
- [15] M. Saleem and M. Usman Akram, "Detection of hemorrhages for diagnosis of malarial retinopathy," *Cairo International Biomedical Engineering Conference*; 11-13 Dec 2014; Giza, Egypt.
- [16] T. Spencer, J.A. Olson, K.C. McHardy, P.F. Sharp and J.V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Computers and Biomedical Research*, vol. 29, 284-302, 1996.
- [17] Y.G. Jiang, C.W. Ngo and J. Yang, "Towards optimal bag-of-features for object categorization and semantic video retrieval," *Conference on Image and Video Retrieval*; 09-11 July 2007; Amsterdam, Netherlands: pp. 494-501.
- [18] S. Lazebnik, C. Schmid and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* New York, USA, vol. 2, pp. 2169-2178, 2006.
- [19] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* Washington, DC, USA, vol. 2, pp.524-531, 2006.
- [20] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," *IEEE 2003 International Conference on Computer Vision*; 13-16 Oct 2003; Washington, DC, USA.
- [21] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," *IEEE International Conference on Computer Vision (ICCV)*; 06-13 Nov 2011; Washington, DC, USA: pp. 2564-2571.
- [22] H. Bay, A. Ess, T. Tuytelaars and L.V. Gool, "Speed-up Robust Features (SURF)," *Computer Vision and Image Understanding* vol. 110, pp. 346-359, 2008.
- [23] T. Güneş and E. Polat, "Yüz ifade analizinde öz nitelik seçimi ve çoklu SVM sınıflandırıcılarına etkisi," *Gazi Üniv. Müh. Mim. Fak. Der.*, vol. 24, pp. 7-14, 2009.
- [24] C. Deng, H. Xiaoferi and J. Han, "SRDA: An efficient algorithm for large-scale discriminant analysis," *IEEE Transactions on Knowledge and Data Engineering* vol. 20, pp. 1-12, 2008.
- [25] Y. Pan, J. Wu, H. Huang and J. Liu, "Spectral regression discriminant analysis for hyperspectral image classification," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, 25 Aug-01 Sep 2012.
- [26] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the 14th international joint conference on Artificial intelligence*, vol. 2, pp. 1137-1143, 1995.
- [27] A. Baratloo, M. Hosseini, A. Negida and G.E. Ashal, "Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity," *Emerg (Tehran)*, vol. 3, pp. 48-49, 2015.