Classification of Retinal Image for Automatic Cataract Detection

Meimei Yang[1], Ji-Jiang YANG[2,3], Qinyan Zhang[1], Yu Niu[2,3], Jianqiang Li[4,]
1. Automation School, Beijing University of Posts and Telecommunications, Beijing, 100876
2. Research Institute of Information Technology, Tsinghua University, Beijing, 100084
3. TNList, Tsinghua University, Beijing, 100084
4. School of Software Engineering, Beijing University of Technology, Beijing, 100022
yangmeimeitongxin@gmail.com, yangjijiang@tsinghua.edu.cn,
zh_qinyan@163.com, niuyu@tsinghua.edu.cn, lijianqiang@tsinghua.org.cn

***Abstract*—Cataract is one of the most common diseases that might cause blindness. Previous research shows that cataract occupies almost 50% in severe visually impairments. Considering the fact that retinal image is one of the most important medical references that help to diagnose the cataract, this paper proposes to use a neural network classifier for automatic cataract detection based on the classification of retinal images. The classifier building procedure includes three parts: pre-processing, feature extraction, and classifier construction. In the pre-processing part, an improved Top-bottom hat transformation is proposed to enhance the contrast between the foreground and the object, and a trilateral filter is used to decrease the noise in the image. According to the analysis of pre-processed image, the luminance and texture message of the image are extracted as classification features. The classifier is constructed by back propagation (BP) neural network which has two layers. Based on the clearness degree of the retinal image, the patients' cataracts are classified into normal, mild, medium or severe ones. The initial evaluation results illustrate the effectiveness of our proposed approach, which has great potential to improve diagnosis efficiency of the ophthalmologist and reduce the physical and economic burden of the patients and society.**

Keywords—retinal image processing; cataract; improved Top-bottom hat transformation; trilateral filter; neural network classifier

I. Introduction

Blindness prevention is an important challenge all over the world. Cataract, glaucoma and age-related macular degeneration are the three biggest diseases to cause blindness [1]. Vision 2020 [1] pointed that there are about 285 million visually impaired people worldwide in 2011. Only cataract occupies almost 50% in visually impaired. Nevertheless, preventable cause is as high as 80% of the total global visual impairment burden. So improving the eye care service especially the pre-detection is of great importance.

Retinal images are widely used by ophthalmologists and primary care physicians for screening of epidemic eye diseases, because they are direct and no harm [2]. From the retinal image, an ophthalmologist can observe (detect) whether the patient has the eye diseases such as cataract [3], glaucoma [4], and diabetic retinopathy [5]. At the same time, changes in the retinal vascular such as vessel width, tortuosity, and branching angle are important indicators for predicting hypertension and cardiovascular diseases [6]. Based on the clearness degree of the retina image, ophthalmologist can find out the current condition of the patient and predict the visual acuity after phacoemulsification [7].

As far as we know, this paper for the first time proposes a BP neutral network classifier for the automatic cataract detection based on the retinal images classification. The classifying procedure mainly includes pre-processing, feature extraction and classification three parts. The pre-processing can improve the quality of the retinal image, and make it better for subsequent processes. In this paper, it uses an improved Top-bottom hat transformation and a trilateral filter [8] to enhance the quality of the retinal image. Feature extraction is necessary before classification because it can discriminate different classes. Based on the experimental comparison, we selects luminance and the texture message of the image as the features for the classifier building. Finally, the neutral network classifier is constructed. Based on the clearness degree of the retinal images, the patients' cataracts are classified into normal, mild, medium or severe ones. The method proved in this paper has great potential to improve the efficiency of the ophthalmologist to diagnose and reduce the physical and economic burden of the patients and society.

The rest of this paper is organized as follows: Section 2 describes related work. Section 3, 4, and 5 are for terminology, implementation, and conclusion respectively.

II. Related Work

Yitao Liang [9] pointed that the retinal image of G channel in RGB color space has the most obvious contrast among the three channels. Weixing Wang [10] used Top-bottom hat transformation with gray-level to reduce the noise and augment the foreground in blood vessel. Top-bottom hat transformation is an effective method to enhance the contrast of the image. Tomasi [11] proposed a bilateral filter which can preserve the edge and reduce noise. While Garnett [8] pointed that the bilateral filter ignore the difference between the edge and the impulse noise of random values. So he designed a trilateral filter to solve the problem.

Siddalingaswamy [12] searched for the optic disc in the retina applying the method of repeated threshold segmentation

This research is supported by the National Science & Technology Pillar Program of China (Grant No.2013BAH19F01 and Grant No.2013BAH19F04).
Corresponding Author: Ji-Jiang YANG, yangjijiang@tsinghua.edu.cn

978-1-4673-5801-9/13/$26.00

based on its luminance. Sekhar [13] successfully detected the brightest region which is the optic disc in the retinal image by morphological method in the retinal image. Texture feature is the internal message of the image and it provides very valuable information. Haralick [14] used the texture feature for image classification. It succeeds in classifying for photomicrographs, aerial photographs, and satellite images. Xiangrong Zhang [15] extracted gray co-occurrence matrix and gray-gradient co-occurrence matrix from the synthetic aperture radar image as a part of the features for classification.

Michael [16] applied the classifier to the retina. The retinal images were classified into normal and abnormal ones. Then the abnormal ones were given to the doctor. This greatly improved doctors' efficiency, but it only cared the abnormal such as exudes. David [17] classified the diabetic retinal images into normal, mild, medium and severe ones by the Learning Vector Quantization network and back propagation network. The parameters are the ratio of vein and the artery and the position of the exudate. But the parameters are not the sole diagnostic criteria for the diabetes. Only when the diabetes is very serious, the retina has something wrong.

The criteria of cataract classification mainly contain three kinds. They are the OXFORD clinical cataract classification and grading system [3], the America Cooperative Cataract Research Group (CCRG) method [18], and the method proposed by Japanese cooperative cataract epidemiology study group [19]. They are all based on slit lamp or the epidemiology method to classify the cataract. When there aren't plenty of ophthalmologists, these methods are invalid.

III. TERMINOLOGY

A. Cataract

The retina mainly contains three parts: optic disc, macular and blood vessel. Fig.1 shows a picture of retina. Optic disc is the highest round disc in the middle of the retina and connects with brain nerves, which reflect the condition of the brain, even full body. Macular is the optical center of the eye and it is high in lutein which makes it appear yellow. The blood vessel is divided into artery and vein. All the blood vessels converge at the optic disc. The vein is darker and wider than the artery. Artery and vein are called big vessel which is defined in this paper. There are some capillaries which attach to the big vessel. They are called small vessels. In this paper, cataract is classified into four different grades: normal, mild, medium and severe. If the optic disc, big vessel and small vessel are all clear, the retina image is a normal one. If only small vessel isn't clear, it is defined as a mild cataract. If the small vessel and big vessel aren't clear, it is defined as a medium cataract. If the optic disc isn't clear, it is defined as a severe cataract. Fig.2 shows the typical image of each category.

B. Retinal image database

The retinal image database contains 504 pictures. In the database the number of normal retina is 367, the mild is 79, the medium is 37, and the severe is 21. All the retinal images are got from the patients with cataract. They are classified by the professional ophthalmologist. The size of the picture is partly 1924*1556 pixels (88 images) and the other is 3888*2592 pixels. Their format is JPEG. They are all RGB color images.

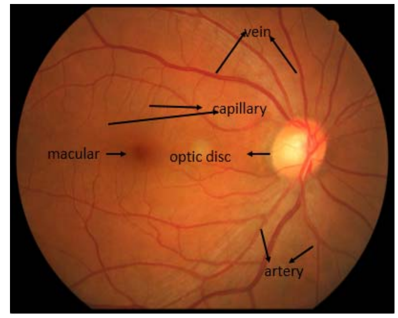


Fig. 1. A picture of retina

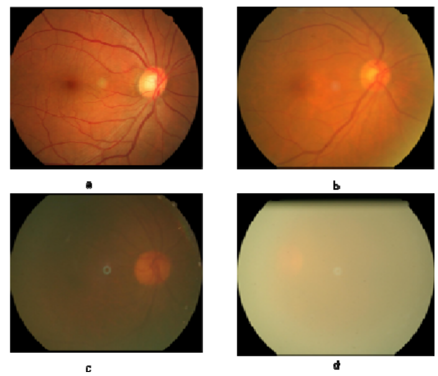


Fig. 2. The typical image of each category (*a is normal, b is mild, c is medium, and d is the severedegree of cataract*)

IV. IMPLEMENTATION

A. Flow chart of classification

Fig.3 shows the flow chart of classification. It mainly contains three parts which are pre-processing, feature extraction, and classifier. Noise can be systematically introduced into images during acquisition and transmission [11]. So in the process of pre-processing, it also needs some methods to improve the quality of the images such as image enhancement and noise reduction. To avoid the problem of the dimension disaster, feature extraction is necessary. Feature extraction using some typical data which is far less than the whole message of the image to represent the whole image. Through studying, classifier can automatically classify the objects into known categories. Through a series of procedures in the flow chart of classification, the retina can be classified automatically.

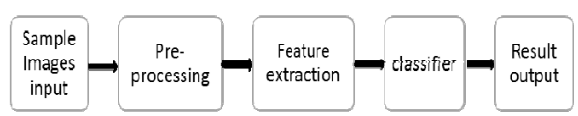
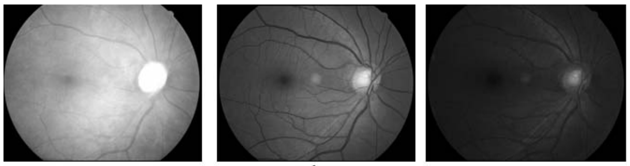


Fig. 3. Flow chart of the classification

B. *Pre-processing*

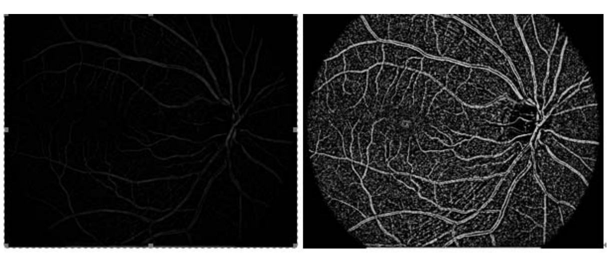
In the step of pre-processing, some personal message in the original image such as name and age was wiped off. Original retinal image is converted to G channel in RGB color space, because it has the most obvious contrast between object and background [9], so it is selected for the next analysis. Fig.4 shows different channel images in RGB color space.



a b c

Fig. 4. Images in R, G, and B channel in RGB color space (*a is for R channel, b is for G channel, and c is for B channel*)

1) Improved Top-bottom hat transformation. To further enhance the contrast between the foreground and blood vessel, improved Top-bottom hat transformation is used. Erosion and dilation are fundmental to morphologic processing. Here denotes A is a picture and B is the structural element in the space of two dimentions (Z^2), which means A and B are all in the same panel. The erosion A by B ($A \ominus B$) is the set of all points in B, which shifts in the panel, and is contained in A. B^* is obtained by the reflection of B about its origin and shifting in the panel. The dilation of A by B ($A \oplus B$) is the set of A contains at least one element in B^*. The image first goes through the erosion and then dilation, then the opening image is obtained. The image first goes through the dilation and then erosion, the closing image is obtained. Top hat transformation means using the original image to substract the opening image. And the Bottom hat transformation means using the original image to substract the closing image. Finally the Top-bottom hat transformation is got when using the Top hat transformation result to substract the Bottom hat transformation result. Top-bottom hat transformation [10] is a mathematic morphologic method that can enhance the quality of the image. But in the retina image, the performance of Top-bottom hat transformation isn't very good according to the result of the experiment. Adding histogram equalization operation after both Top and Bottom hat transformation, the image becomes clearer. This method is the improved Top-bottom hat transformation. The structural element is a disk of 20 pixels in this paper. Fig.5 shows the results of normal Top-bottom and our improved Top-bottom hat transformation. It shows that the blood vessel message becomes more obvious. However, there appears a lot of noise in the image at the same time.



a b

Fig. 5. Images after Top-bottom hat transformation and improved Top-bottom hat transformation (*a is the result of normal Top-bottom hat transformation; b is after the result of improved Top-bottom hat transformation*)

2) Trilateral filter. Trilateral filter is used to remove the noise in the image after the improved Top-bottom hat transformation.

Tomasi [11] pointed that the difference between pixel at edge and pixel of impulse noise: at least one neighbor pixel has the same gray value with the pixel at the edge, while the pixel of impulse noise isn't. Denote x is the central pixel in consideration and y is a neighbor pixel of x. The weighting function of each y with respect to x in bilateral filter $\omega(x, y)$ is defined as equation (1):

$$\omega(x,y) = \omega_S(x,y)\omega_R(x,y) \tag{1}$$

where

$$\omega_S(x,y) = e^{-\left(\frac{x-y}{2\delta_s}\right)^2} \tag{2}$$

and

$$\omega_R(x,y) = e^{-\left(\frac{\mu_x-\mu_y}{2\delta_r}\right)^2} \tag{3}$$

The ω_S weighting function decreases as the spatial distance between x and y increasing, and the ω_R weighting function decreases as the radiometric "distance" between the intensities μ_x and μ_y increasing. The parameters of δ_S and δ_R control the behavior of the weights of ω_S and ω_R respectively. They serve as rough thresholds for identifying pixels sufficiently close spatially or radio metrically. So bilateral filter can decrease noise and preserve the message of edge well. The weights have been normalized in the bilateral filter.

But the bilateral filter can't discriminate the pixel at edge and pixel of impulse noise. A ω_I weighting function is introduced into trilateral filter [8].

$$\omega(x,y) = \omega_S(x,y)\omega_R(x,y)^{1-J(x,y)}\omega_I(x)^{J(x,y)} \tag{4}$$

where

$$J(x,y) = 1 - e^{-(ROAD_m(x)+ROAD_m(y)/2)^2/2\delta_J^2} \tag{5}$$

$$\omega_I(x) = e^{-ROAD_m(x)^2/2\delta_I^2} \tag{6}$$

and

$$ROAD_m(x) = \sum_{i=1}^{m} r_i(x) \tag{7}$$

Equation $r_i(x)$ is the absolute difference between x and one of its neighbor pixels. Variable $ROAD_m(x)$ is sum of the m minimum of $r_i(x)$ (m is a nonnegative integer).

When x is an impulse noise,, $ROAD_m(x)$ is big and $J(x,y) \to 1$, so ω_I comes into action in the trilateral filter. When x is a pixel at edge, $J(x,y) \to 0$, ω_I is almost no effect and it is a bilateral filter. In this way, the filter can discriminate the pixel at the edge and the pixel of impulsive noise. The parameter δ_i (in Equation (6)) controls the behavior of its weights. It serves as a rough threshold for identifying pixels sufficiently close to impulsive noise. This paper uses this filter twice. The parameters in the trilateral are 3, 1, 0.1, 0.1, 0.1 and 6, 20, 0.1, 0.1, 0.1 (they represent the value of $m, \delta_S, \delta_R, \delta_i$ and δ_j respectively). Fig.6 shows the images after trilateral filter. The original images are the images in Fig.2.

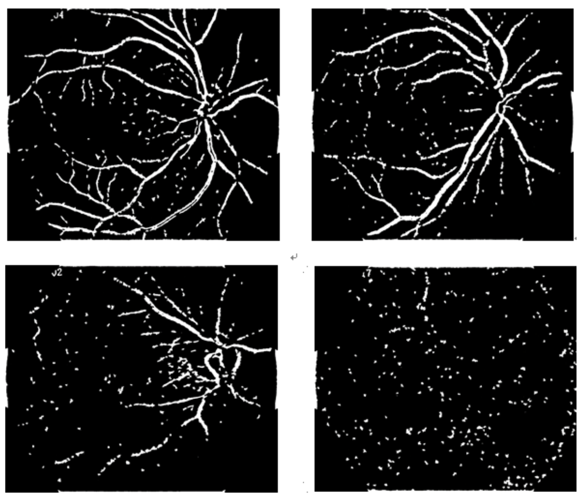


Fig. 6. Images after trilateral filter

C. Feature extraction

To avoid the problem of the dimension disaster, feature extraction is necessary. Feature extraction uses some typical data which is far less than the whole message of the image to represent the whole image. There are 40 features extracted from the image.

1) Lumiance feature. The clearness of the retinal image is the standard for cataract classification, so the lumiance is selected as a feature for classification. In this paper, this feature is extracted from the image which has been wiped off personal message. Then the image is converted to a binary image and the threshold is 0.6. The lumiance feature is how much the white pixel share in the image.

2) Gray co-occurrence matrix. Fig.6 shows that the differences between four classes are very big. So gray co-occurrence matrix is selected to do the job of feature extraction. This matrix provides the gloal message of the image. For example, angular second moment feature is a mesure of homogenity of the image, and contrast is a measure of constrast of the image. From the gray occurrence matrix, we extract 24 features from the degree of 0, 45, 90, and 135 [14]: angular second moment, correlation, entropy, contrast, inverse difference moment and sum of squares. Image after pre-processing is transformed from 256 gray levels to 8 levels. It computes much quicker than the image before transformed [14]. Then these features in gray co-occurrence matrix are extracted.

3) Gray-gradient co-occurrence matrix. The gray-level co-occurrence matrix usually is an essential feature for object identification, while it cannot provide the edge information. Gray-gradient co-occurrence matrix concerns the associated statistic distribution of the gray and the edge gradient. From the gray-gradient co-occurrence matrix, the following 15 features can be computed [15] : little gradient dominance, large gradient dominance, gray heterogeneity, gradient heterogeneity, energy, gray average, gradient average, gray mean square error, gradient mean square error, correlation, gray entropy, hybrid entropy, inertia and inverse difference moment.These features are extracted from the same images with gary co-occurrence.

D. BP neural network classifier.

Appropriate designed classifier can automatically classify the objects into known categories. In this paper, BP neural network is chosen to finish the classification work. We use the pattern recognition toolbox (NPR) in MATLAB to realize this classifier. The BP neural network has two layers, and the number of the hidden neurons is 10 in the paper. The number of neurons in input layer is 40 and the number of neurons in output layer is 4. All the features extracted from the images in the database are input into this classifier.

The data is divided three parts: training data, validation data and test data. The training data accounts for 70% (352 images), and the other two data both take 15% (76 images). The training data decides the weights and bias in the network. The network is trained by the method of conjugate gradient decent. Training automatically stops when its performance stops improving, which is indicated by the increasing of the mean square error of the validation samples. That is to say the validation data are to prevent the network overlearning. The images in test data are used to test the performance of the network in recognition the class of the images whose classes are unknown in advance.

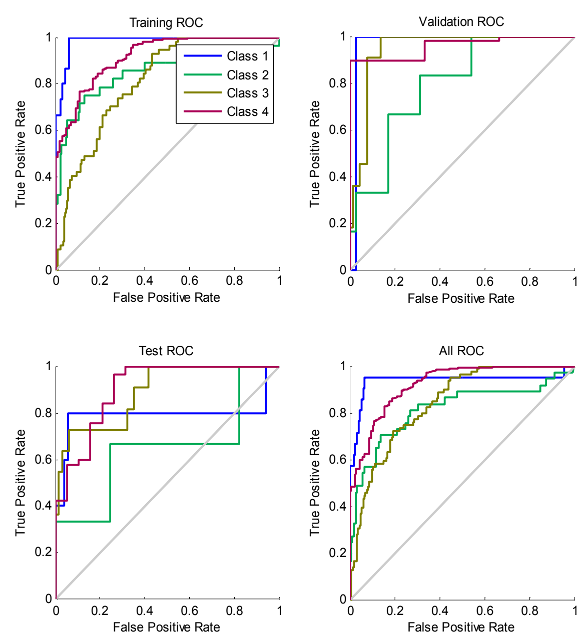


Fig. 7. Receiver operatation characteristic (ROC)

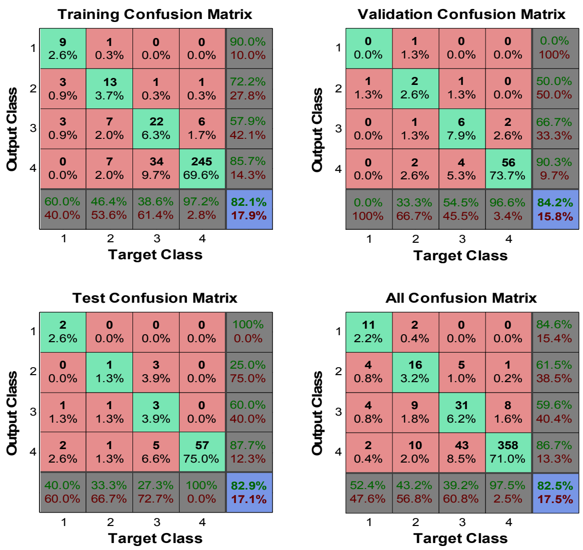


Fig. 8. Confusion matrix

E. The experimental results analysis

There are two main parameters reflecting the performance of the classifier. They are receiver operating characteristic (ROC) and confusion matrix. ROC almost has no relation with the distribution of the images in four classes, while confusion matrix has. Fig.7 is the ROC. The ideal result is the false and true positive rate is 0 and 1 (point *(0,1)*). Class 1, 2, 3 and 4 are for severe, mild, medium and normal cataract respectively. The result of class 4 is better than others in Fig.7.

Fig.8 is the confusion matrix. The horizontal axis is the target class and vertical axis is the output class. *C(m, n)* represents that target class is *m* and output class *n* (*m, n*=1, 2, 3, 4). When *m* equals *n*, the classification is correct. Here, an example is used to introduce the specific meaning of every *C(m, n)* in the training confusion matrix. *C(1,1)* means the classification is correct. In *C(1,1)*, the number 9 represents the number of image in *C(1,1)* and 2.6% is how much the number 9 accounts for in all the training data. The number 82.1% in low right corner is the sum of true positive rate of 4 kinds of classes in the training data. The number 17.9% is the sum of error rate of 4 kinds of classes in the training data. The second number 46.4% in the left bottom represents the true positive rate when target class is class 2. The third number 57.9% in the right row represents the true positive rate when output class is class 3. Fig.8 shows the training and test true positive rate is 82.1% and 82.9% respectively.

The time of pre-processingis about 375 and 1200 seconds for one image whose size is 1924*1556 and 3888*2592 respectively, which is too long. A trilateral filter for binary image will be designed in the future to solve this problem.The time for feature extraction and for NPR is less than 5 minutes for 504 retinal images, and it is acceptable.

Comparably, methods proposed in [3], [18] and [19], our method can automatically classify the degree of cataract. This cataract classification criterion comes from reference [7], whose authors (also our cooperators) are the experts of ophthalmology and their research comes from the real clinic practice and statistics. We believe our work is the first to realize the automatic cataract retinal image classification through IT technology. Hitherto, there is no similar work for comparison in literatures. In next step, we will continue following this criterion and collect extra retina images for larger-scale trials. And we hope our work could provide a baseline for the research work in this field.

V. CONCLUSION

A neural network classifier is proposed to automatically classify the severity of cataract. It is based on the clearness degree of the retinal image. The classifier consists of pre-processing, feature extraction and classification three parts. An improved Top-bottom hat transformation and trilateral filter are used to improve the quality of the image. The features are extracted from luminance and texture. The classifier is a two layer BP neural network. Through the classifier, the patients' cataracts are classified into normal, mild, medium or severe ones. It has great potential to improve the efficiency of the ophthalmologist and help to reduce the physical and economic burden of the patients and society as well.

In the future, we will continue our research with the emphasis on extracting the features of blood vessel in retinal image, and reduce the time in pre-processing. Furthermore, running the classification in larger trials will be explored.

ACKONWLEDGMENT

Thank Dr. Xu Liang and MD. Yang Hua from Beijing TONGREN Hospital for providing the database and professional guidance.

REFERENCES

[1] http://www.iapb.org/vision-2020.

[2] Bhuiyan, A., Lamoureux, E., Nath, B., Ramamohanarao, K., & Wong, T. Y., "Retinal image matching using hierarchical vascular features", Computational intelligence and neuroscience, Year 2011.

[3] Bron, A. J., Brown, N. A. P., Ayliffe, W., & Hill, A. R., "The Oxford clinical cataract classification and grading system", International ophthalmology, Vol.9, No.4, pp.207-225, Year 1986.

[4] P. Mitchell, H. Leung, J.J. Wang, E. Rochtchina, A.J. Lee, T.Y. Wong, &R. Klein, "Retinal vessel diameter and open-angle glaucoma: the Blue Mountains eye study", Ophthalmology,Retinal vessel diameter and open-angle glaucoma: the Blue Mountains eye study, Vol.112, No.2, pp.245–250, Year 2005.

[5] Clara I.Sánchez, María García, Agustín Mayo, María I.López, Roberto Hornero, "Retinal image analysis based on mixture models to detect hard exudates", Medical Image Analysis, Vol.13, No.4, pp.650–658, Year 2009.

[6] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, & B. van Ginneken, "Ridge based vessel segmentation in color images of the retina", IEEE Trans. Medical Imaging, Vol.23, No.4, pp.501–50, Year 2004.

[7] XU Liang, YANG Chuan-wu, YANG Hua, WANG Shuang, SHI Yu-ying, & SONG Xu-dong, "The study of predicting the visual acuity after phacoemulsification according to the blur level of fundus photography", Ophthalmol CHN, Vol.19, No.1, pp.46-50, Year 2010.

[8] Garnett, R., Huegerich, T., Chui, C., & He, W., "A universal noise removal algorithm with an impulse detector", Image Processing, Vol.14, No.11, pp.1747-1754, Year 2005.

[9] Yitao Liang, Lianlian He, Chao Fan, Feng Wang, & Wei Li, “Preprocessing Study of Retinal Image Based on Component Extraction”, Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education, pp.670-672, Year 2008.

[10] Weixing Wang, Hao Song, & Qi Zhao, “A modified Watersheds Image Segmentation Algorithm for Blood Cell”, International Conference on Communications, Circuits and Systems Proceedings, Vol.1, pp.450- 454, Year 2006.

[11] Tomasi, Carlo, & Roberto Manduchi. "Bilateral filtering for gray and color images", Computer Vision, Sixth International Conference on IEEE, pp.839-846, Year 1998.

[12] Siddalingaswamy P.C., & G. K. Prabhu, “Automated Detection of Anatomical Structures in Retinal Images”, International Conference on Computational Intelligence and Multimedia Applications, Vol.3, pp.164-168, Year 2007.

[13] S. Sekhar, W. Al-Nuaimy, & A. K. Nandi, “Automated localization of retinal optic disk using hough transform”, 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp.1577-1580, Year 2008.

[14] Haralick R M, Shanmugam K, & Dinstein I H, “Textural features for image classification”, IEEE Transactions on Systems, Man and Cybernetics, Vol.6, pp.610-621, Year 1973.

[15] Xiangrong Zhang, Tan Shan, & Licheng Jiao, “SAR Image Classification Based on Immune Clonal Feature Selection”, Springer-Verlag Berlin Heidelberg, pp.504–511, Year 2004

[16] Michael Goldbaum, Saied Moeui, Adam Taylol, Shankar Chatterjee, Jeff Boyd, Edward Huntel, & Ramesh Jain, “Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images”, International Conference on Image Processing, Vol.3, pp.695-698, Year 1996.

[17] J. David, Rekha Krishnan, & Sukesh Kumar, “A Neural Network Based Retinal Image Anonent Extraction”, Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education, Year 2008.

[18] L.T. Chylack Jr, M.R. Lee, W.H. Tung, & H.M. Cheng, “Classification of human senile cataractous changes by the American Cooperative Cataract Research Group (CCRG) method”, Investigative ophthalmology & visual science, Vol.24, No.4, pp.424-431, Year 1983.

[19] Sasaki, K., Shibata T., Obazawa H., & Fujiwara T., "Classification system for cataracts", Ophthalmic research, Vol.22, No.1, pp.46-50, Year 1990.