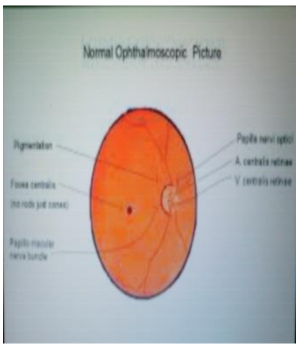
Detection of Cataracts via Classification

Irrum Shafique, Ayesha Fatima, Aliya Zafar
Department of Computer Engineering, College of EME, National University of Sciences and Technology (NUST), H-12, Islamabad, Pakistan
Irumshafique07@gmail.com, ayeshaf1339@gmail.com, Aliya.zafar24@gmail.com

***Abstract*—Lens opacity classification system whose objective is to minimize the leading cause of blindness which is cataract which results from opacification of lens. Current methods for the diagnosis of cataracts are partially based on grading protocol and needs considerable manual intervention which are not very much prone to subjective analysis. In this work a system is proposed to automatically classify cataracts and non cataracts images. Both gray level and color based features were implemented. The proposed algorithm was tested by images from a community study. The success rate of classification of cataracts is 91.8% for 41 images. Dataset was randomly divided into 10 folds.1 fold was used for testing and rest of the 9 folds were used for training. Classification accuracy of 91.8% has been achieved. The experimental results shows that the proposed approach is promising in clinical diagnosis.**

1. INTRODUCTION

Cataract is common disease that is caused due to pollution in environment, certain intakes and age related factors in humans are also responsible for cloudiness of eye retina. It is one of the main cause of blindness in need aware or need blind countries in terms of medical diagnostics or accessibility of medical aid in remote areas [1].cataract is a progressive disease, but early detection and diagnosis of cataracts can save vision. Recent innovations proved that one out of five patients with different cataract symptoms is infected with this disease. The common symptoms of cataracts are blur and faded vision, glare, poor night vision otherwise, people suffering from diabetes, smokers and alcohol addicts are continuously at risk of getting cataract[2].a healthy optic lens consist of aqueous humor, cornea, pupil, dynamic crystalline lens and retina.In location based categories of cataracts i.e. nuclear cataracts (NC) commonly known as age-related or "senile" cataracts whose major cause is aging process grows from center of the retina lens and distant vision is affected in this way. Cortical cataract expands itself in the periphery of the lens effects inward in a spoke-like manner, its main cause is diabetes that's why often called diabetic cataracts. Posterior subcapsular cataract (PSC) Progresses through back of the lens. it is found in patients mostly with background of diabetes, steroid and mainly severe nearsightedness. Cataract is a progressive disease in which it decreases ones visual capabilities, lack of accurate equipment cheap and massive eye health diagnostics. Functioning of eye is similar to camera, when light passes through eye passing and then enters cornea and forwarded via aqueous humor, pupil and then into lens. Lens bends in order to focus specific object on optic lining on the back side of eye. Image passing through cells of retina and optic nerve are retrieved back to brain finally, which process those images. Cataracts forms where protein is stored in cells and its buildup in the lens that makes it cloudy.



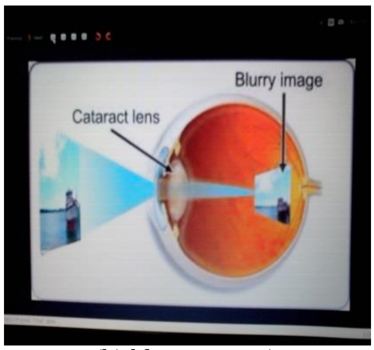


Fig.1 ((a) normal lens (b) blurry image)

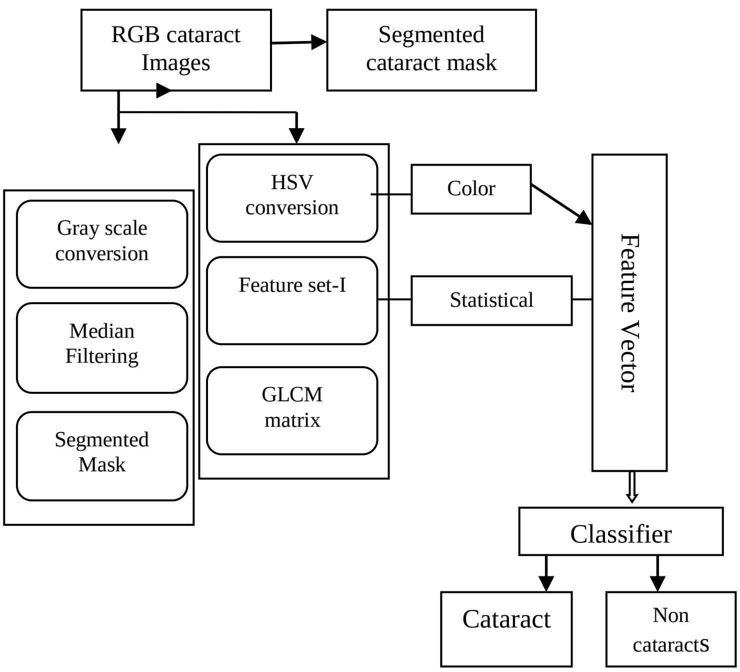
Formation of cataracts creates hindrance of light from passing through the lens. It causes considerable loss of vision. Therefore newly regenerated cells on the external side of the lens, previous cells are centered aligned with center of the lens resulting in the cataract [3]. Major causes due to which cataract progresses are aging, enzymes, heredity issues, diabetes, renal failure. Whereas daily routine habits effects more than external pollution factors e.g. drugs consumption, alcohol UV consumption, excessive use of microwave or infrared radiations and trauma etc can also cause cataracts. An eye exam will be given to test how well you can see? An analysis test with or without your glasses should be taken physician will ask you to dilate your pupils for medical examination of condition of the lens and other parts of the eye. Purpose behind this project is to analyze eye lens images for reliable detection of proteins present in lens external wall and finding considerable region of interest.

2. Previous Work

A number of automated system for optic lens clarity, vision enhancement, detection of region of interest, and specific patterns and circle in lens image have been proposed. An automated feature based analysis of slit lamp images of mouse lens has been proposed. A standard contrast enhancement technique was applied by [3] for identification of genotype of the slit lamp images based upon automated computer analysis of lens structural phenotype.In this system neural network was discrimination of genotype/phenotype in which 85% genotype of slit lamp images were classified correctly. For grading of nuclear cataracts support vector machine(SVM) is used in order to train the model to do grade prediction. The technique proposed by [4] is based upon local features and similarity based landmarks. Nuclear cataracts usually classified upon the severity of opacity using slit lamp images. In this system anatomical structure in lens is detected via active shape model (ASM). Through this technique more than 5000 images were tested and for feature extraction success rate were 95% and grading error was 0.36. This automated analysis can help to improve grading objectivity. Xu et al. [5] propose a system to automate the grading for measuring severity of nuclear cataract from slit lamp images via sparse group regression. This selected feature will comprehend selection of other secondary features their parameters and their regression model for training it collectively. On experimental basis 5378 images were used in order to train the model and this system outperform with absolute error of .336%. Li et al.[6] presented the detection of region of interest in order to measure the opacity of cortex of the lens so that severity of cataract should be graded. a spoke like feature is extracted which subtract the desired region from other opacities. Overall this system is tested upon 611 samples from which 466 were used for testing and their mean error of opacity area is 3.15% and exact cortical grading was 85.6%.Li et al.[6] have used edge detection in order to do pupil detection for detecting strongest edges on convex hull or ellipse fitting using non linear least square. Further, this opacity is graded via radial edge detection and further post processed to discriminate amongst other opacities. For auto grading results are further processed using Wisconsin grading protocol. After application of this experimental model 98% accuracy is achieved and minimum error was 7% while, grading severity of cortical 86.3 were graded correctly.

3. Methodology

a. Workflow Diagram



b. ROI detection (pupil localization)

The preprocessing step of our algorithm is to extract region of interest (pupil) in input images because maximum information is available. Here assumption we made that pupil is circular in shape hence we applied Hough transform because of its effectiveness towards discontinuous pixels in our dataset so that it can identify arbitrary shapes. In this process we used an intensity image and then converted t o binary to extract region of interest. In our work auto pupil detection is combination of edge detection, Hough transform in order to localize pupil.

In preprocessing we used RGB images and converted RGB image into gray scale and then we apply median filter on it. To remove an artifacts error and for better result of classification we are using median filter here because we will use some statistical features like standard deviation, mean, contrast etc which may be sensitive for noise. And finally in preprocessing we apply segmented mask to segment the candidate opacity. So after the preprocessing we have segmented candidate region as an output.

c. Feature Extraction

Feature extraction is the procedure that takes out the hidden properties of the image that is further used in the classifier. In this stage we are doing Feature extraction. In Feature Extraction we are taking some Statistical Features from GLCM and some other features from HSI color space.

d. GLCM for feature extraction

In statistical texture analysis texture features are calculated from the statistical distribution of observed combination of vigor's at specified position relative to each other in image. According to the number of pixels in each amalgamation, statisticsare categorized into first-order, second-order and higher-order statistics. The Gray Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features. For GLCM the most common statistics (Energy, Contrast, Correlation and Homogeneity) are computed from candidate lesion and used as features. In these features. Correlation is a measure of gray level linear dependence between the pixels at the specific location relative to each other.Also one can notice that cataract images have more contrast than other region as a lot of having opacity with in dark regions so we take contrast feature. We made a feature set from these four features and called feature set1.

After extraction of these four features we trained our classifier and test the accuracy which was 83% at that time so we think of adding some more features in order to achieve more accuracy, so we collect some new features as describe below.

e. HSV color space features:-

After less accuracy with only included the features of GLCM, we then think to include some color information and then collect the new features from color information made a new feature set called feature set2. The color examination can be performed in different color spaces, each one with different properties. RGB is the most popular color space. But it has a series of disadvantages: it is not perceptually uniform and it depends on the acquisition setup and shows correlation between the three color channels. There are some alternatives like Hue saturation and brightness (HSV) color spaces which perform explanation of color similarly to human one that are perceptually uniform color spaces.

In HSV, the Hue (H) of a color refers the resemblance of color, the Saturation(S) tells us the impurity of colors and the Value of a color also called its lightness; describe how dark the color is. We then convert RGB cataract image into (HSV) color space. Now in HSV color space to include color information we calculate Mean Hue, Mean Saturation, Standard Deviation (SD) Hue, SD Saturation, SD intensity. We also observe in our database images some images have less energy in infected region then other mostly belong to cataract class, so we also calculate the median intensity of opacity region. Due to the reason of non uniform illumination we didn't rely on just energy we also calculate the gradient of opacification region with its border region.

After calculating these features now we have some statistical information features and some features with color information. From these features we made a feature vector. And similarly after the features calculation of all images of our database of

41 images we store all this information in the form of database.

4. Classification of Cataracts

For Classification we test four different types of Classifier to test which classifier result are more accurate in this dataset of 41 images. Four different Classifiers include; Support Vector Machine (SVM), Naïve Bayes, K Nearest Neighbor (KNN), and Decision Tree. In KNN classification is done by a reference set which contains both targeted and input variable and comparing the unknown which contain only the input variables to the reference set. The distance of the unknown to the k nearest neighbours finds its class assignment by either averaging the class numbers of k nearest points. Naive Bayes assumes that presence or absence of a particular feature is unrelated to the presence or absence of any other feature, given the class variable. Decision Tree classifier poses a series of carefully crafted questions about the attributes of each record. Each time it receives a answer, a follow up question is asked until a conclusion about the class label of the record is established,

For classification we use 10 Folds to check the result. Before passing all the observation to classifier we randomly distribute the 41 observations of all images of both Cataract and Non-cataract. Then we perform classification with all the classifier and noticed that the accuracy of KNN classifier is better among all of them with the average accuracy of 91.8%.

5.Results

The detailed results of all classifier's performance are shown in Fig (2).

Classifier	Time	Accuracy	Sensitivity	SD Sensitivity	Specifity	SD Specifity	Conf Matrix	Average Accuracy
KNN	0.0094 0.0031 0.0134	93 91.5 91	77.50 75 70	3.8188	96.88 95.63 96.25	0.6250	155 5 9 31	91.8
SVM	0.0043 0.0031 0.004	86.5 87.5 88	42.50 45.00 47.00	2.2546	97.50 97.50 97.88	0. 2194	150 10 17 23	87.3
Naïve Bayes	0.0030 0.0018 0.0027	87 86.5 88	62.50 60.00 65.00	2.5	93.13 93.13 93.75	0. 3580	151 9 15 25	87.1
Classification Tree	0.0047 0.0039 0.0036	86 88 87	75.00 67.50 65.00	5.2042	88.75 93.13 92.50	2.3680	146 14 16 24	86.3
Average Feature Calculation Time			1.2990 sec					

6.Conclusion

In this document, we have mentioned the activities of different classifier as analytic resources to aid doctors in the detection of cataract even at the beginning.These classifiers are also appropriate for identifying the effectiveness of cataract functions using the post-catact pictures.

However, these resources usually do not generate outcomes with 100% precision. The precision of these resources rely on several aspects such as the dimension and high company's training set, the rigor of the coaching imparted, and parameter chosen to signify the feedback. From the outcomes acquired as it is obvious that the KNN classifier is efficient to the track of about 91.8% precision.

7. REFERENCES

[1]. Yuen J, Li Y, Shapiro LG, et al. Automated, Computerized, Feature-Based Phenotype Analysis of Slit Lamp Images of the Mouse Lens. Experimental eye research 2008;86(4):562-575. doi:10.1016/j.exer.2007.11.019.

[2]. Huiqi Li; Joo Hwee Lim; Jiang Liu; Mitchell, P.; Tan, A.G.; Jie Jin Wang; Tien Yin Wong, "A Computer-Aided Diagnosis System of Nuclear Cataract,"Biomedical Engineering, IEEE

Transactions on , vol.57, no.7, pp.1690,1698, July 2010
[3]. Huiqi Li; Liling Ko,; Joo Hwee Lim; Jiang Liu; Wong, D.W.K.; Tien Yin Wong, "Image based diagnosis of cortical cataract," Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE , vol., no., pp.3904,3907, 20-25 Aug. 2008
[4]. Huiqi Li; Liling Ko; Joo Hwee Lim; Jiang Liu; Wong, D.W.K.; Tien Yin Wong; Ying Sun, "Automatic opacity detection in retro-illumination images for cortical cataract diagnosis," Multimedia and Expo, 2008 IEEE International Conference on , vol., no., pp.553,556, June 23 2008-April 26 2008
[5]. Jagadish Nayak Automated Classification of Normal, Cataract and Post Cataract Optical Eye Images using SVM Classifier
[6]Xu, Y., Gao, X., Lin, S., Wong, D. W. K., Liu, J. et al. Automatic grading of nuclear cataracts from slit-lamp lens images using group sparsity regression. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, Lecture Notes in Computer Science 8150