Identifying Abnormalities in the Retinal Images using SVM Classifiers

Shantala Giraddi
BVB College Of Engg.and Technology Dept of CSE, Hubli, India

Jagadeesh Pujari, Ph.D
SDM College Of Engg.and Technology Dept of ISE, Dharwad, India

Shivanand Seeri
BVB College Of Engg.and Technology Dept of CSE, Hubli, India

ABSTRACT
Automated early detection of exudates in retinal images is a challenging task. With the global diabetic population increasing at an alarming rate, there is need for development of automated systems for detection of exudates. The main obstacle in exudates detection is extreme variability of color and contrast in retinal images that depends on the degree of pigmentation, size of the pupil and illumination. The aim of this paper is to develop and validate systems for detection of hard exudates and classify the input image as normal or diseased one. The authors have proposed and implemented novel method based on color and texture features. Performance analysis of SVM and KNN classifiers is presented. Images classified by these classifiers are validated by expert opthamalagists. .

Keywords
Diabetic Retinopathy, SVM classifier, KNN classifier Exudates.

1. INTRODUCTION
Diabetes mellitus (DM) is a major cause of blindness all over the world. It has been found that patients with diabetic retinopathy (DR) are 25 times more likely to become blind than non-diabetics [1]. There were 31.7 million diabetics in India in year 2000 which is expected to reach 79.4 million by year 2030[2]. Both type 1 and type 2 diabetes can cause Diabetic Retinopathy. Both eyes can be affected by Diabetic retinopathy. Often there are no early signs of diabetic retinopathy. Symptoms may only become noticeable once the disease advances. Typical symptoms of retinopathy are [3].

- Dark strings floating in your vision (floaters)
- Blurred vision
- Fluctuating vision
- Dark or empty areas in your vision
- Vision loss
- Difficulty with color perception

Diabetic retinopathy may be classified as early diabetic retinopathy and advanced diabetic retinopathy. Early diabetic retinopathy is called non proliferative diabetic retinopathy (NPDR). In this stage , there are no major symptoms. Retinal swelling may be present to some extent. In this stage tine capillaries become semi permeable membranes, later leaking or oozing fluid and blood into the retina. As the disease progress, the smaller vessels may close and the larger retinal vessels may begin to dilate and become irregular in diameter. Advanced diabetic retinopathy also called Proliferative diabetic retinopathy (PDR) is the most severe type of diabetic retinopathy. At this stage, damaged blood vessels begin to break, leak blood into the clear, jelly-like substance that fills the center of eye. They are not able to supply the nutrients to the retina. The starvation of nutrients in the retina causes growth of new blood vessels. This growth of new capillaries is called neo vascularization.

Diabetic patients who have had diabetes for more than five years are likely to develop some form of Diabetic Retinopathy. Only regular screening can result in early detection and effective management of DR. Patients should get their both eyes screened at least once in a year. These screening programs generate large number of images and processing of which is time consuming process. Automated diabetic Retinopathy can save time and reduce the workload of opthmalagists. Developing strategies for screening large population for early detection of DR is engaging attention of several groups in India. Automated grading is less costly and of similar effectiveness, it is likely to be considered a cost- effective alternative to manual grading.

In several patients, the only visible symptoms of DR are Exudates [4]. Hard exudates occurring in the macula can cause significant visual impairment. The main obstacle in exudates detection is extreme variability of color and contrast in retinal images that depends on the degree of pigmentation, size of the pupil and illumination. These factors affect the appearance of exudates in the retinal images. Many techniques such as clustering, morphological operations, pixel wise classification using various classifiers like Back propagation neural network (BPNN), Support vector machines (SVM) have been employed for the exudates detection. All these techniques have high computational requirement.

In their previous work, authors have proposed and validated novel method for exudate detection based on color, texture features using Back propagation neural network [5]. In this paper comparative analysis of SVM and KNN classifiers is presented. The rest of the paper is structured as follows. In section 2, related works are discussed. Section 3 deals with the proposed methodology. In Section 4 experimental results are explained in detail. Paper is concluded in section 5

2. RELATED WORK
Hussain F.Jaafar et.al [6] proposed a method based on top-down image segmentation and local thresholding by a combination of edge detection and region growing. Grading of hard exudates is performed. Nidhal K et.al [7] proposed a system in which Exudates are found by thresholding and false exudates are separated from true exudates using first order texture features. Ratio of red to green channel is used to identify optic disc region. Xiwei Zhang et.al [8] Segmented exudates based on mathematical morphology and characterized candidates based on classical features as well as contextual features and validated their system on e-ophtha EX database. In Another approach proposed by Juan Martin Cardenas [9] et.al image preprocessing is done with mean shift filtering and region growing algorithm is performed from local maxima regions taken as seeds to get final results. M. Usman Akram et.al [10] used filter banks to extract the

candidate regions for possible exudates. Spurious exudates regions are eliminated by removing the optic disc region. Then Bayesian classifier is applied to detect exudates and nonexudate regions. Vimala et.al [11] proposed a method using LAB color space image. The preprocessed color retinal images are segmented using Line operator and Fuzzy C Means Clustering technique in order to detect Optic Disk. The exudates are extracted using K-means clustering and finally the classification is done using SVM. Zeljkovic et.al [12] proposed the automated algorithm that applies mathematical modeling which enables light intensity levels emphasis, easier exudates detection, efficient and correct classification of retina images. According to authors the proposed algorithm is robust to various appearance changes of retinal fundus images. Shraddha Tripathi et.al [13] proposed a system based on Differential Morphological Profile (DMP) using green channel of the image which is thresholded to get a binary image. Optic disc is eliminated based on area and shape index which is nearly 1 for circular objects

3. MATERIALS AND METHODS

This section presents the proposed methodology for the classification of retinal image into healthy image or the diseased one using SVM classifier and KNN classifier. The proposed methodology is given in Fig 1. The input image is divided into blocks as shown in the Fig 3. From each block we extract various color features and texture features. Authors have considered two models for their study. Model I consists of first order color features like mean, variance, skewness and kurtosis from red and green channel of the image. Performance of SVM and KNN classifiers is compared. Model II consisted of second order texture features based on GLCM like homogeneity, contrast, correlation and variance simultaneously. These features are fed to SVM and KNN classifiers and their performance is compared

3.1 Image Dataset

Publicly available database, DIARETDB0 is considered for the study. The DIARETDB0 database consists of 81 color retinal images; each image is of size 1152X1500 in the png format. The database consists of images of varying color and contrast. Figure 2 shows typical retinal image having Diabetic Retinopathy

3.2 Preprocessing

The input images are color images of size 1150*1500. The input image is scaled to 575X 750. Images are contrast enhanced using adaptive histogram equalization (CLAHE) which operates on small data regions rather than the entire image. For computation of First order features, green channel and red channel of the image are extracted. For extracting second order features, the scaled images are converted into grey scale images.

3.3 Partitioning of ROI

Retinal image has many clinical structures such as optic disc and vessels which pose major obstacles in the detection of exudates. The common approach of extracting features based Gray Level Co-occurrence Matrix (GLCM) by constructing the GLCM over the entire image may not work here since retinal image has many anatomical structures like blood vessel and optic disc. The input image is divided into a number of smaller blocks of sizes 36X37 pixels. Based on its content, the block is categorized. The blocks are numbered 1-320. The creation of the blocks is shown in Fig 3.

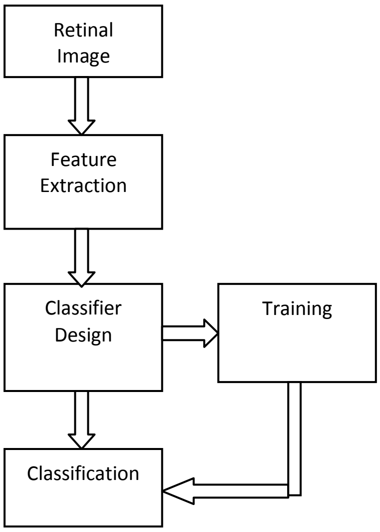

Figure 1. Proposed Methodology

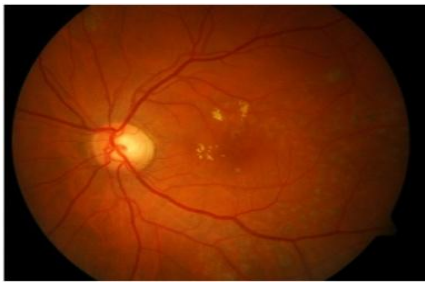

Figure 2: Retinal image with Diabetic retinopathy

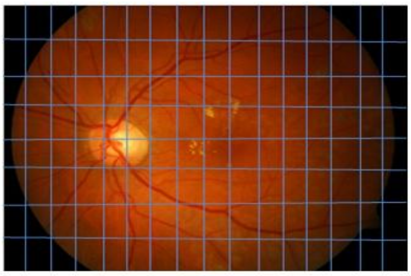

Figure 3: Retinal image divided into blocks

The blocks are divided into four categories namely

1. Normal blocks – Blocks containing Normal Retinal part
2. Exudates blocks – Blocks containing exudates
3. Vessel blocks – Blocks containing vessels
4. Optic disc blocks- Blocks containing part of Optic disc.

From these blocks, first order color features and second order texture features (GLCM) are extracted.

3.4 Feature Extraction

Pixel intensities are simplest available features useful for pattern recognition. Intensity features are first order statistics that depend only on individual pixel values. The intensity and its variation inside the retinal images can be measured by features like: median, mode, standard deviation and variance. Various first and second order features (haralick) based on GLCM extracted from input image are listed below in Table 1.

TABLE 1. First order and second order features extracted

First order features	Second order features
Contrast	Mean
Correlation	Variance
Energy	Skewness
Homogeneity	Kurtosis
	Energy
	Entropy

3.5 SVM Classifiers

SVM is one of the best known methods in pattern recognition and image classification. It is basically a linear classifier but by using different types of kernels one can model non linear classifier. It is designed to separate a set of training images into two different classes,(x1, y1), (x2, y2), ..., (xn, yn) where xi in Rd, d-dimensional feature space, and yi in {-1,+1}, the class label, with i=1..n [1]. SVM builds the optimal separating hyper planes based on a kernel function (K). All images, of which feature vector lies on one side of the hyper plane, belong to class -1 and the others are belong to class +1. The design of the SVM classifier architecture is very simple and mainly requires the choice of the kernel, cost C and gamma. Nevertheless, these have to be chosen carefully since an inappropriate kernel can lead to poor performance. Support Vector Machines are inherently designed for binary classification. For multiclass problems as in image classification, one needs an appropriate multi-class method. The two approaches commonly used,

- One-Against-One (1A1) Combines several binary classifiers: and applies pair wise comparisons between classes.
- One-Against-All (1AA) compares a given class with all the others put together.

The accuracies of both these methods are almost the same [14]. Ultimately the choice of technique adopted is a personal preference and depends on the uniqueness of the dataset at hand. In the "one against the others" algorithm, n SVM models are constructed and i^{th} model is trained with examples in i^{th} class with positive labels and other examples with negative labels. In this way n hyper planes are constructed, where n is the number of classes. Each hyper plane separates one class from the other classes. Comparatively "one against the others" is having lesser complexity and has been chosen for this study.

3.6 KNN classifiers

The k–nearest neighbor classifier is one of the most important conventional nonparametric classifier. It was introduced by Fix and Hodges [15]. It has been proved to be a simple and powerful recognition algorithm that provides good performance for optimal values of k. In the k-nearest neighbor rule, a test sample is assigned the class most frequently represented among the k nearest training samples. If two or more such classes exist, then the test sample is assigned the class with minimum average distance to it.

4. RESULTS AND DISCUSSIONS

In this section, experimental results of classification of retinal image using SVM classifier and KNN classifier is presented. The experiments were conducted on Intel Core 2 Duo 2.66 GHz PC with 3GB of RAM using MATLAB2009.b. A population of 40 images is used for the study. Totally 100 blocks with exudates, 100 normal blocks, 100 blocks with vessels and 100 blocks with optic disc are used for training. 20 images, 10 images with exudates and 10 healthy images are used for testing. Each image is divided into 320 blocks of equal size. The image having exudates of more than 2 blocks is classified as abnormal image. Otherwise image is classified as normal image. The ground data is verified by expert opthamalagist.

4.1 Experiment I: SVM Classifier

Experiments were conducted with SVM-RBF using color and texture features. With color features, C value maintained at a constant value of 0.8 and for texture features C value maintained at 1. For each value of C, Gamma value is varied (0.4, 0.6, 0.8 and 1). The classifier performance with color features for different value of gamma is shown in the Fig 4. Fig 5 shows SVM performance with texture features. It can be observed that with color features, Gamma parameter does not much impact on %tp which is 83.4% .With GLCM features higher value of gamma leads increased %tp(80% for G=1.2) and lesser %fp (8.62% for G=1.2). The best classification accuracy was obtained for Gamma value of 1.2.

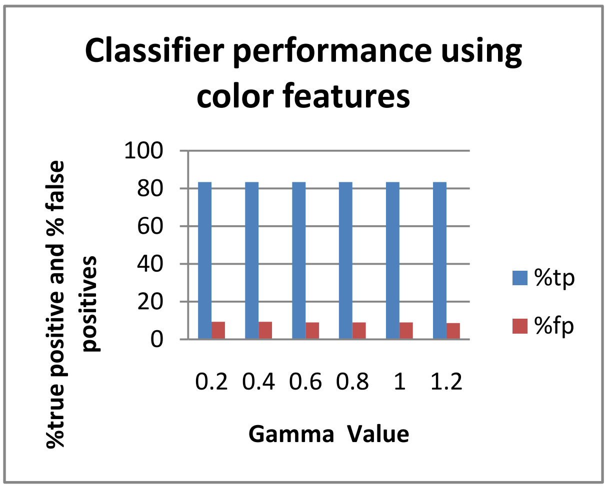

Figure 4. Classification accuracy of SVM with different value of Gamma (color features)

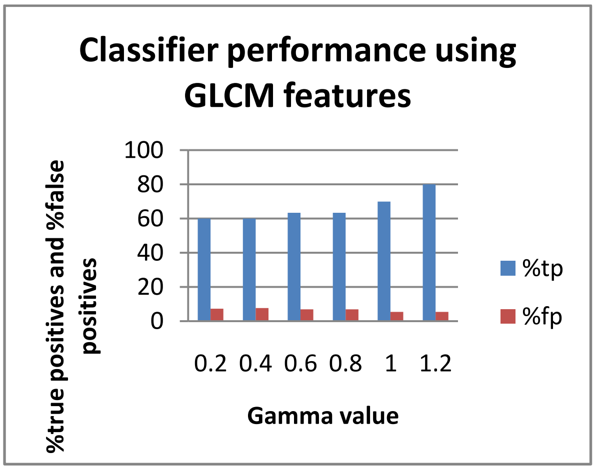

Figure 5. Classification accuracy of SVM with different value of Gamma (texture features)

4.2 EXPERIMENT II: KNN CLASSIFIER.

Experiments were conducted with KNN using color and texture features with various values of K. Different distance measures like city block, Euclidian, cosine distances can be used. Euclidian distance, which is most commonly used distance metric, has been used in the study. Fig 6 shows the classification accuracy with different values of K with color features. Highest tp of 92% is obtained with k value of 8 (corresponding %fp is 12.18%) and a lowest tp of 70% value is obtained with a k value of 2(corresponding %fp is 11.56%).

Fig 7 shows classifier performance with GLCM features. Highest %fp obtained with k value 8(% fp 85% corresponding %to is 39%) and lowest %fp obtained with k=1, 2, 3 (65%) and corresponding %fp is 23%.

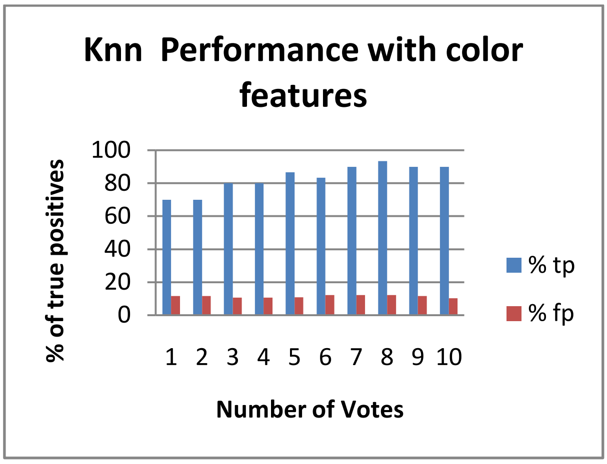

Figure 6. Classification accuracy of Knn classifier using color features with different number of votes

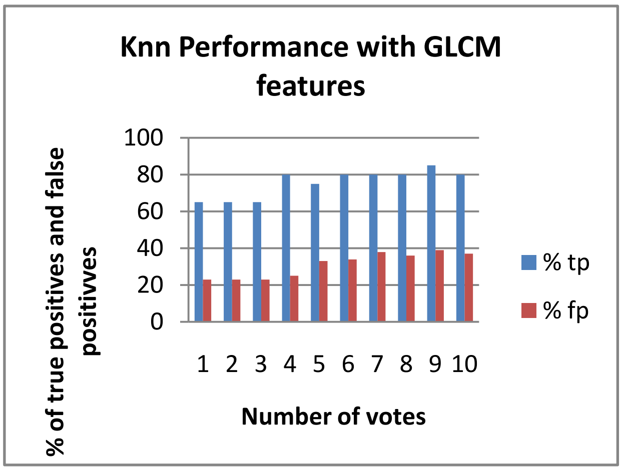

Figure 7. Classification accuracy of Knn classifier using texture features with different number of votes

In case of KNN classifier, increasing the number of votes has increased the % of true positive, optimal value of K being around 8 or 9 (with color features as well as texture features). With SVM classifier using color features, the value of gamma did not have any effect on result whereas SVM using texture features, increasing the gamma value increased the %tp and decreased %fp.

5. CONCLUSION

The results demonstrate that color features are good in identifying exudates than texture features however texture features based on GLCM may help in reducing the number of false positives. The %tp with SVM classifier is 83.4 and %tp with Knn classifier is 92%. Knn classifier outperforms SVM with color as well as texture features. False exudates cannot be eliminated completely. In future work, authors plan to focus reducing the number of false positives.

6. REFERENCES

[1] Abha Gupta Caroline Styles Brtitish Journal Of Diabetes and Vascular Disease doi: 0.1177/1474651410374192 British Journal of Diabetes & Vascular Disease September/October 2010 vol. 10 no. 5 224-230

[2] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004; 27: 1047

[3] http://www.mayoclinic.org/diseases- conditions/diabetic-retinopathy/basics//symptoms/con-20023311 20023311

[4] C.I.Sanchez Roberto Hornero , Mar´ıa I. L´opez , Mateo Aboy ,Jes´us Poza , Daniel Ab´asolo " A Novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis" ,Medical Engineering and Physics , pp 350-357, Issue 30 2008.

[5] Shantala Giraddi, Dr Jagadeesh Pujari, "Role of GLCM features in Identifying Abnormalities in the Retinal Images", Accepted IJGSP To be published

[6] H.F. Jaafar, A.K. Nandi and W. Al-Nauimy, "Automated detection of exudates in retinal images using a split-and-merge algorithm," EUSIPCO 2010, Alborg, pp. 1622-1626, 2010.

[7] Nidhal K. El Abbadi Enas Hamood Al- Saadi , "Automatic Detection of Exudates in Retinal Images", International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013

[8] H.F. Jaafar, A.K. Nandi and W. Al-Nauimy, "Automated detection of exudates in retinal images using a split-and-merge algorithm," EUSIPCO 2010, Alborg, pp. 1622-1626, 2010.

[9] Nidhal K. El Abbadi Enas Hamood Al- Saadi , "Automatic Detection of Exudates in Retinal Images", International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013

[10] Xiwei Zhanga, , , Guillaume Thibaulta, Etienne Decencièrea, , Beatriz Marcoteguia, Bruno Laÿd, Ronan Dannod, Guy Cazuguele, f, Gwénolé Quellecf, Mathieu Lamarde, f, Pascale Massinb,

[11] Agnès Chabouisc, Zeynep Victorb, Ali Erginayb "Exudate detection in color retinal images for mass screening of diabetic retinopathy" DOI: 10.1016/j.media.2014.05.004

[12] Juan Martin Cardenas, M.Elena Martinez-Perez ,Francesc March, Nidiyare Hevea-Montiel "Mean Shift Based Automatic Detection Of Exudates in retinal Images" Advances in Intelligent Systems and Computing Volume 184, 2013, pp 73-82.

[13] Gidudu Anthony, Hulley Gregg and Marwala Tshilidzi

[14] Image Classification Using SVMs: One-against-One Vs One-against-All, http://arxiv.org/ftp/arxiv/papers/0711/0711.2914.pdf

[15] Fix, E. and HODGES, J. L. JR. (1951). Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties. Report No. 4, Project No. 21-49-004, USAF School of Aviation Medicine.