International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:8, No:5, 2014

Comparing SVM and Naïve Bayes Classifier for Automatic Microaneurysm Detections

A. Sopharak, B. Uyyanonvara, S. Barman

***Abstract*—**Diabetic retinopathy is characterized by the development of retinal microaneurysms. The damage can be prevented if disease is treated in its early stages. In this paper, we are comparing Support Vector Machine (SVM) and Naïve Bayes (NB) classifiers for automatic microaneurysm detection in images acquired through non-dilated pupils. The Nearest Neighbor classifier is used as a baseline for comparison. Detected microaneurysms are validated with expert ophthalmologists' hand-drawn ground-truths. The sensitivity, specificity, precision and accuracy of each method are also compared.

***Keywords*—**Diabetic retinopathy, microaneurysm, Naïve Bayes classifier, SVM classifier.

I. Introduction

MICROANEURYSMS (MAs) are the earliest clinically characteristic of diabetic retinopathy (DR) [1], [2]. Retinal MAs are focal dilatations of retinal capillaries. They appear as small round dark red dots on the retinal surface. The diameter of a MA lies between 10 and 100μm [2], [3]. Early screening for DR could improve the prognosis and prevent the blindness. Currently, manual examination by ophthalmologists is given but the process takes time and the number of ophthalmologists is not sufficient to cope with all patients. The automatic MA detection could help enable early detection of DR and could help ophthalmologists track the process of treatment over time.

Several methods have been reported in literatures. A few attempts are based on morphological technique. T. Spencer et al. [3], M. J. Cree et al. [4] and A. Frame et al. [5] employ a mathematical morphology technique to segment MA within fluorescein angiograms. T. Walter et al. [6] propose a method based on diameter closing and kernel density estimation for automatic classification. Some are based on region growing. C. Sinthanayothin et al. [7] propose an automated system of detection of diabetic retinopathy using recursive region growing segmentation (RRGS). D. Usher et al. [8] employ a combination of RRGS and adaptive intensity thresholding to detect candidate lesion regions and a neural network is used for classification. Clustering has also been proposed. B. Dupas et al. [9] use a diameter-closing to segment MA candidate regions and k-nearest neighbours (kNN) to classify MA. M. Niemeijer et al. [10] combine prior works by T. Spencer et al. [3] and A. Frame et al. [5] with a detection system based on pixel classification and new features are proposed. A kNN classifier was used in the final step. B. Zhang et al. [11] use multi-scale correlation coefficients (MSCF). They detect coarse MA candidate using MSCF and fine MA using features classification.

Few attempts are based on SVM. B. F. Zohra and B. Mohamed [12] proposed a computer-based system to identify normal, NPDR and PDR using the SVM. X. Wen-Hua [13] proposed a detection of MAs in bifrequency space based on SVM.

In previous work, we have proposed MA detection using naïve Bayes classifier [14]. Here SVM classification is applied. The results are compared with kNN classifier.

The paper is organized as follows: image preprocessing, feature extraction, naïve bays classifier, SVM classifier and nearest neighbor classifier are proposed in Section II. In Section III, the results of all methods are presented and compared. The conclusion is given in Section IV. The SVM classifier is a substantial improvement on previous work.

II. Method

All digital retinal images are taken without pupil dilation with a KOWA-7 non-mydriatic retinal camera with a 45° field of view and a size of 752 x 500 pixels. Noise removal and contrast enhancement are applied to the original images in pre-processing step. A shade correction algorithm is applied to the green band in order to remove slow background variation due to non-uniform illumination. The optic disc and vessels are then removed using mathematical morphology [15].

A. Naïve Bayes Classifier

A total feature set contains 18 features are proposed to distinguish MA pixels from non-MA pixels. The list of 18 features is shown in Table I.

Sopharak. A. is with the Faculty of Science and Arts, Burapha University Chanthaburi Campus, 57 Moo.1 Kamong, Thamai, Chanthaburi, Thailand 22170 (phone: +66-39-310000; fax: +66-39-310128; e-mail: akara@buu.ac.th).

Uyyanonvara. B. is with the Sirindhorn International Institute of Technology, Thammasart University, and 131 Moo. 5 Tiwanont Road, Bangkadi, Muang, Pathumthani, Thailand (e-mail: bunyarit@siit.tu.ac.th).

Barman. S. is with the Kingston University, Penrhyn Road, Kingston Upon Thames, Surrey, KT1 2EE, United Kingdom (e-mail: S.Barman@kingston.ac.uk).

TABLE I
MA EXTRACTED FEATURES

No.	Description
1	*The pixel's intensity value of shade corrected image (I_{sc})*
2	*The pixel's intensity value of green band image after preprocessing (I_g)*
3	*The pixel's hue*
4	*The standard deviation of shade corrected image.* A window size of 15x15 is used.
5	*The standard deviation of green band image after preprocessing.* A window size of 15x15 is used.
6-11	*Six Difference of Gaussian (DoG) filter responses. The DoG filter subtracts one blurred version of an original image from another blurred version of the image [16]. We convolve with seven different Gaussian kernels with standard deviations of 0.5, 1, 2, 4, 8, 16, and 32. We use DoG1, DoG2, DoG3, DoG4, DoG5 and DoG6 to refer to the features obtained by subtracting the image at scale σ = 0.5 from scale σ = 1, scale σ = 1 from σ = 2, scale σ = 2 from σ = 4, scale σ = 4 from σ = 8, scale σ = 8 from σ = 16, and scale σ = 16 from σ = 32, respectively*
12	*The area of the candidate MA*
13	*The perimeter of the candidate MA*
14	*The eccentricity of the candidate MA*
15	*The circularity of the candidate MA*
16	*The mean intensity of the candidate MA on shade corrected image*
17	*The mean intensity of the candidate MA on green band image*
18	*The ratio of the major axis length and minor length of the candidate MA*

The naive Bayes classifier [16]-[18] uses the principle of Bayesian maximum a posteriori (MAP) classification: measure a finite set of features $\mathbf{x} = (x_1,\ldots, x_n)$ then select the class

$$\hat{y} = \arg\max_y P(y|\mathbf{x})$$

where

$$P(y|\mathbf{x}) \propto P(\mathbf{x}|y)P(y) \quad (1)$$

$P(\mathbf{x}|y)$ is the likelihood of feature vector $\mathbf{x}$ given class y, and $P(y)$ is the priori probability of class y. Naive Bayes assumes that the features are conditionally independent given the class:

$$P(\mathbf{x}|y) = \prod_i P(x_i|y)$$

We estimate the parameters $P(x_i|y)$ and $P(y)$ from training data.

We used Weka data mining software [19] for feature discretization and naive Bayesian classification.

We first estimate the model of (1) from a training set using all features then evaluate the resulting classifier's performance on a separate test set. Then we iteratively delete features until the average of the precision (PR) stops improving. On each step, for each feature, we delete that feature from the model, train a new classifier, and evaluate its performance on the test set. The PR of the best such classifier is compared to the PR of the classifier without deleted features. If PR improves, we permanently delete that feature then repeat the process. Finally, the best feature set and classifier are retained.

B. Support Vector Machines Classifier

A support vector machine [20], [21] is originally defined for the classification of linearly separable classes of objects based on the statistical learning theory. SVM can also used to separate nonlinear classes by mapping the coordinate of the objects into a feature space using nonlinear function. The nonlinear mapping induces by the feature function is computed with kernel function. SVM will separate the data with a hyperplane that leaves the maximum margin between two classes.

SVM map training data into a high-dimensional feature space in which we can construct a separating hyperplane maximizing the margin, or distance from the hyperplane to the nearest training data points.

In the input space, a binary SVM's decision function can be written:

$$\hat{y} = h(\mathrm{x}) = sign(\sum_{i=1}^{n} \alpha_i y_i K(\mathrm{x}, \mathrm{x}_i) + b) \quad (2)$$

where x is the feature vector to be classified, i indexes the training examples, n is the number of training examples, y_i is the label (1 or -1) of training example i, $K(\cdot,\cdot)$ is the kernel function, and α_i and b are fit to the data to maximize the margin. Training vectors for which $\alpha_i \neq 0$ are called support vectors.

We used libSVM's [20] implementation of the ν-SVM with the radial basis function kernel on a 20-node Gnu/Linux Xeon cluster for training and testing SVM classifiers. For a given feature set, to find optimal hyperparameters (ν, the tolerance for misclassified training examples, and γ, the width of the radial basis function) for the SVM, we performed a grid search, retaining the parameter values for which test set accuracy is maximized.

The best feature set obtained from naive Bayes are used as an initial feature set for the SVMs. We then add features to the SVMs classifier one at a time and compare the PR of each classifier to that of the previous classifier. The first feature added in is always the last feature removed during the naive Bayes classifier's feature selection process. The feature-adding process is repeated until all features are added back. The best feature set is the set which provides the highest PR.

C. Nearest Neighbor Classifier

Nearest neighbor classifier with Euclidean and Mahalanobis distance metrics are used as our baseline for comparison. To be able to compare with naive Bayesian and SVM classifiers, we used the best feature sets obtained for naive Bayesian and the SVM.

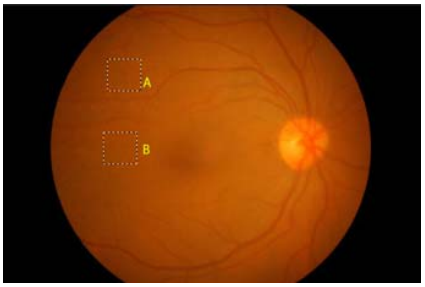
III. Results

Naive Bayesian is tested on Weka data mining software running on standard PC while SVMs and nearest neighbor are tested on a 20-node Gnu/Linux Xeon cluster. Finally, detected MAs are compared with the ophthalmologists' hand-drawn ground-truth images for verification. We fit the naive Bayesian model to the training set using all 18 features. We removed features from the classifier one by one and compared the resulting PR to PR obtained on the previous feature set. We continued this process until the PR stopped improving.

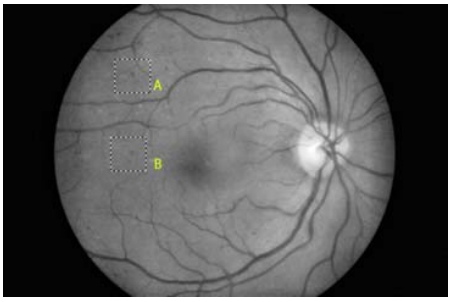
Finally, the best features for the naive Bayesian contained 10 features: 1.the pixel's intensity of shade corrected image, 2. the pixel hue, 3. the standard deviation of shade corrected image, 4. DoG4, 5. the area of the candidate MA, 6. the perimeter of the candidate MA, 7. the eccentricity of the candidate MA, 8. the circularity of the candidate MA, 9. the mean intensity of the candidate MA on shade corrected image and 10. the ratio of the major axis length and minor length of the candidate MA. The overall sensitivity, specificity, precision, and accuracy are 84.82%, 99.99%, 89.01%, and 99.99%, respectively.

For the SVM, the best performance is obtained using 12 features: 1.the pixel's intensity of shade corrected image, 2. the pixel hue, 3. the standard deviation of shade corrected image, 4. DoG4, 5. the area of the candidate MA, 6. the perimeter of the candidate MA, 7. the eccentricity of the candidate MA, 8. the circularity of the candidate MA, 9. the mean intensity of the candidate MA on shade corrected image, 10. the ratio of the major axis length and minor length of the candidate MA, 11. DoG3 and 12. DoG6 with $\nu = 0.004$ and $\gamma = 0.995$. The overall sensitivity, specificity, precision, and accuracy are 85.82%, 99.99%, 92.02%, and 99.99%, respectively.

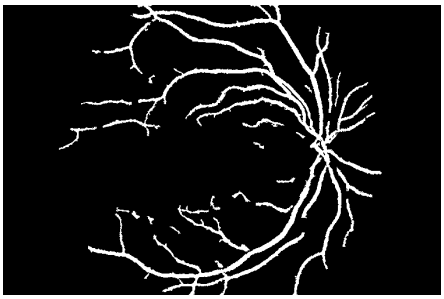
On best feature set obtained from the naive Bayesian classifier, the nearest neighbor classifiers (Euclidean and Mahalanobis distances) have a precision of 75.01% and 75.71%, respectively. On the best feature set obtained from the SVM classifier, the nearest neighbor classifier achieved a precision of 78.15% and 79.09%, respectively. The results indicate that the naive Bayesian and SVM classifiers perform substantially better in PR than the nearest neighbor classifier. In addition, the nearest neighbour classifier using the best feature set obtained from the SVM classifier performs better than that using the best feature set for the naive Bayesian classifier. Example image of diabetic retinopathy retinal image and shade corrected image are shown in Fig. 1. Result images of MAs detection from naïve Bayes classification and SVM classification are shown and compared in Fig. 2.



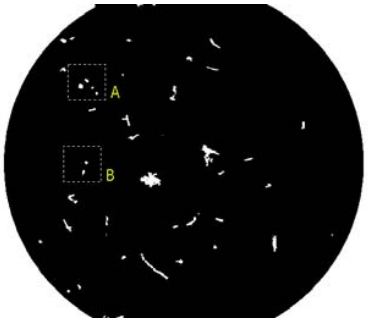
(a)



(b)



(c)



(d)

Fig. 1 Retinal image with MAs cropped (Sample A and Sample B). (a) Original image in RGB (b) Shade corrected image (c) Detected vessels image (d) Candidate MA detection

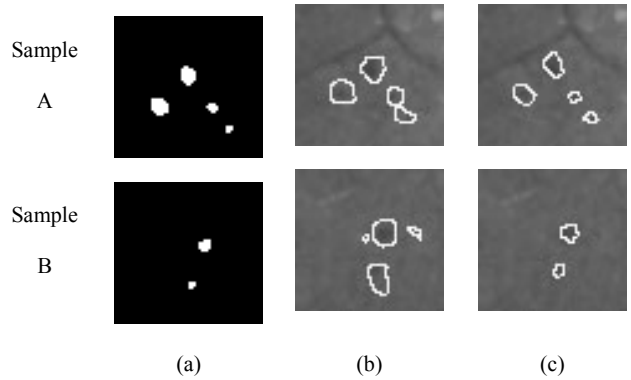


Fig. 2 Results of MAs detection (a) Ground-truth images (b) Naive Bayes classification results (c) SVM classification results

IV. Conclusion

In this paper we propose applied SVM to MAs detection and compared the result with previous work of MAs detection with naive Bayesian classifier and kNN classifier. SVM classifier performs the best precision. In future work, we plan to explore using the system as a practical aid to help ophthalmologists for diabetic retinopathy screening.

Acknowledgment

This research is funded by the Burapha University, Chanthaburi Campus and National Research University Project of Thailand Office of Higher Education Commission (Thammasat University).

References

[1] S. Wild, G. Roglic, A. Green et al., "Global prevalence of diabetes: estimates for the year 2000 and projections for 2030," *Diabetes Care* 27, 2004, pp.1047-1053.

[2] P. Massin, A. Erginay, and A. Gaudric, "Retinopathie Diabetique", Elsevier, Editions scientifiques of medicales, Elsevier, SAS, Paris 2000.

[3] T. Spencer, J.A. Olson, K.C. McHardy et al., "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comp Biomed Res* 29, 1996, pp. 284–302.

[4] M.J. Cree, J.A. Olson, K.C. McHardy et al., "A fully automated comparative microaneurysm digital detection system," *Eye* 11, 1997, pp. 622–628.

[5] A. Frame, P. Undrill, M. Cree et al., "A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms," *Comput. Biol. Med.* 28, 1998, pp. 225–238.

[6] T. Walter, P. Massin, A. Erginay et al., "Automatic detection of microaneurysms in color fundus images," *Medical Image Analysis* 11(6), 2007, pp.555-566.

[7] C. Sinthanayothin, J.F. Boyce, T.H Williamson, T.H. et al., "Automated Detection of Diabetic Retinopathy on Digital Fundus Image," *Diabetic Medicine* 19(2), 2002, pp. 105–112, 2002.

[8] D. Usher, M. Dumskyj, M. Himaga et al., "Automated Detection of Diabetic Retinopathy in Digital Retinal Images: A Tool for Diabetic Retinopathy Screening," *Diabetic Medicine* 21(1), 2004, pp. 84–90.

[9] B. Dupas, T. Walter, A. Erginay et al., "Evaluation of automated fundus photograph analysis algorithms for detecting microaneurysms, haemorrhages and exudates, and of a computer-assisted diagnostic system for grading diabetic retinopathy," *Diabetes & Metabolism* 36(3), 2010, pp. 213-220.

[10] M. Niemeijer, B. van Ginneken, J. Staal et al., "Automatic detection of red lesions in digital color fundus photographs," *IEEE Trans Med Imaging* 24(5), 2005, pp.584-592.

[11] B. Zhang, X. Wu, J. You et al., "Detection of microaneurysms using multi-scale correlation coefficients," *Pattern Recognition* 43(6), 2010, pp. 2237-2248.

[12] B. F. Zohra and B. Mohamed, "Automated diagnosis of retinal images using SVM", Faculte des Science, Department of Informatique, USTO, Algerie.

[13] X. Wen-Hua. "Detection of microaneurysms in bifrequency space based on SVM." *Electronics, Communications and Control (ICECC), 2011 International Conference on*. IEEE, pp. 1432-1435, 2011.

[14] A. Sopharak, B. Uyyanonvara, and S. Barman., "Automatic Microaneurysm Quantification for Diabetic Retinopathy Screening." *International Conference on Image Analysis and Processing*, pp. 2591-2594, 2013.

[15] A. Sopharak, B. Uyyanonvara, S. Barman et al., "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Computer Medical Imaging and Graphics* 32(8), 2008, pp. 720-727.

[16] R.S. Kenneth, C.R. John, J. Matthew et al. (2006, June 15). Difference of Gaussians Edge Enhancement [Online]. Available: http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/diffgaussians/index.html

[17] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*. Vol. 29, pp.131-163, 1997.

[18] O.D. Richard, E.H. Peter, and G.S. David, *Pattern Classification 2nd edition,* A Wiley-Interscience Publication, 2000, pp. 20-83.

[19] X.Y. Wang, J. Garibaldi, and T. Ozen, "Application of The Fuzzy C-Means clustering Method on the Analysis of non Pre-processed FTIR Data for Cancer Diagnosis," *Internat. Conf. on Australian and New Zealand Intelligent Information Systems (ANZIIS)*, pp. 233-238, 2003.

[20] C.C. Change and C.J Lin. (2001). LIBSVM: A library for support vector machines [Online]. Available: http://www.csie.ntu.edu.tw/cjlin/libsvm

[21] P.H. Chen, C.J. Lin and B. Scholkopf, "A Tutorial on v-Support Vector Machines,: *Applied Stochastic Models in Business and Industry* 21(2), 2005, pp. 111-136.