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

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

where

\[
P(\mathbf{x} | y) \propto P(\mathbf{x} | y) P(y)
\]

\( P(\mathbf{x} | y) \) is the likelihood of feature vector \( \mathbf{x} \) given class \( y \), and \( P(y) \) is the prior 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 be used to separate nonlinear classes by mapping the coordinate of the objects into a feature space using nonlinear function. The nonlinear 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:

\[
\hat{y} = h(\mathbf{x}) = \text{sign} \left( \sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \right)
\]

where \( \mathbf{x} \) 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 \( \alpha_i \) 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 \( \nu \)-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 (\( \nu \), the tolerance for misclassified training examples, and \( \gamma \), 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

Naïve 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 naïve 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 naïve 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 naïve 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 naïve 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 naïve 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.
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


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.