

Machine learning approach to automatic exudate detection in retinal images from diabetic patients

Akara Sopharak^{a*}, Matthew N. Dailey^b, Bunyarit Uyyanonvara^a
Sarah Barman^c, Tom Williamson^d, Khine Thet Nwe^b and Yin Aye Moe^b

^a*Sirindhorn International Institute of Technology, Thammasat University, Thailand*; ^b*Computer Science and Information Management, Asian Institute of Technology, Thailand*; ^c*Digital Imaging Research Centre, Kingston University, UK*; ^d*St Thomas' Hospital, UK*

(Received 29 October 2008; final version received 15 June 2009)

Exudates are among the preliminary signs of diabetic retinopathy, a major cause of vision loss in diabetic patients. Early detection of exudates could improve patients' chances to avoid blindness. In this paper, we present a series of experiments on feature selection and exudates classification using naive Bayes and support vector machine (SVM) classifiers. We first fit the naive Bayes model to a training set consisting of 15 features extracted from each of 115,867 positive examples of exudate pixels and an equal number of negative examples. We then perform feature selection on the naive Bayes model, repeatedly removing features from the classifier, one by one, until classification performance stops improving. To find the best SVM, we begin with the best feature set from the naive Bayes classifier, and repeatedly add the previously-removed features to the classifier. For each combination of features, we perform a grid search to determine the best combination of hyperparameters ν (tolerance for training errors) and γ (radial basis function width). We compare the best naive Bayes and SVM classifiers to a baseline nearest neighbor (NN) classifier using the best feature sets from both classifiers. We find that the naive Bayes and SVM classifiers perform better than the NN classifier. The overall best sensitivity, specificity, precision, and accuracy are 92.28%, 98.52%, 53.05%, and 98.41%, respectively.

Keywords: exudate; diabetic retinopathy; naive Bayes classifier; support vector machine; nearest neighbor classifier

1. Introduction

Diabetic retinopathy is a severe eye disease and a major cause of blindness. Exudates, lipid leakages from blood vessels, are visible signs of an early stage of retinal abnormality in diabetic retinopathy. Diabetic patients need regular screening because early detection of exudates could help prevent blindness. However, manual examination by ophthalmologists takes time and the number of experts is not sufficient to meet the demand for screening. Given the limitations of manual screening, the prospect of automatic detection of retinal exudates, towards diagnosis and tracking

the progress of a patient's treatment program, is enticing.

There have been several attempts to solve this problem. Quite a few are based on thresholding and region growing:

- Liu et al. (1) detect exudates using thresholding and region growing. Their fundus photographs were taken with a nonmydriatic fundus camera then scanned by a flatbed scanner.
- Ege et al. (2) use a median filter to remove noise, segment bright lesions and dark lesions by thresholding, perform region growing, then identify exudates

This is a preprint of a paper published in the Journal of Modern Optics. The paper should be cited as "Sopharak, Akara, Dailey, Matthew N., Uyyanonvara, Bunyarit, Barman, Sarah, Williamson, Tom, Nwe, Khine Thet and Moe, Yin Aye (2010) 'Machine learning approach to automatic exudate detection in retinal images from diabetic patients', Journal of Modern Optics, 57(2):124-135"

*Corresponding author. Email: aj.akara@gmail.com

regions with Bayesian, Mahalanobis, and nearest neighbor (NN) classifiers. The system failed to detect exudates in low quality images.

- Sinthanayothin et al. (3) report the result of an automated detection of diabetic retinopathy on digital fundus images using recursive region growing segmentation (RRGS). They measure performance on 10×10 pixel patches.
- Usher et al. (4) detect candidate exudate regions using a combination of RRGS and adaptive intensity thresholding. The candidate regions thus extracted are classified as exudate or non-exudate by a neural network.
- Kavitha and Shenbaga (5) propose median filtering and morphological operations for blood vessel detection. They use multilevel thresholding to extract bright regions assumed to be the optic disc or exudates. They detect the optic disc as the converging point of the blood vessels, then classify the other bright regions as exudates. The method performed poorly on low-contrast images.

Thresholding and region growing methods are straightforward, but selecting threshold values, region seed points, and stopping criteria are difficult.

Clustering has also been proposed as a possible solution to the exudate detection problem:

- Osareh and colleagues (6, 7) use fuzzy *c*-means clustering to segment color retinal images into homogeneous regions, then train neural networks and support vector machines (SVMs) to separate exudate and non-exudate areas.

- Zhang et al. (8) use local contrast enhancement and fuzzy *c*-means clustering in the LUV color space to segment candidate bright lesion areas. They use hierarchical SVMs to classify bright non-lesion areas, exudates, and cotton wool spots.

The main difficulty with clustering methods is determining the number of clusters to use.

A few other attempts are based on specialized features and morphological reconstruction techniques:

- Katarzyna et al. (9) detect candidate exudate regions with a watershed transformation and a marker and extract the optic disc based on geodesic reconstruction by dilation.
- Sanchez et al. (10) combine color and sharp edge features to detect exudate. First they find yellowish objects, then they find sharp edges using various rotated versions of Kirsch masks on the green component of the original image. Yellowish objects with sharp edges are classified as exudates.
- Walter et al. (11) use morphological reconstruction techniques to detect contours typical of exudates.
- Wang et al. (12) extract color features then use a nearest neighbor classifier to identify retinal lesions.

All of these techniques are highly sensitive to image contrast.

Most of the work just reviewed is based on imagery acquired after dilating patients' pupils, e.g. with eye drops, to make exudates and other retinal features clearly visible. Since pupil dilation takes time and is uncomfortable for patients, in our work, we investigate methods for automated exudate detection on imagery acquired without pupil dilation. The ultimate aim of our work is to develop an automatic

exudate detection system to provide decision support and reduce ophthalmologists' workloads.

In previous work, we have proposed and evaluated methods for automatic detection of exudate in non-dilated retinal images using mathematical morphology techniques (13), fuzzy *c*-means (14), and a combination of fuzzy *c*-means and mathematical morphology (15). In experiments on comparable data sets, the sensitivity and specificity for these methods were 80.00% and 99.46%, 92.18% and 91.52%, and 86.03% and 99.36%, respectively. While these results are encouraging, they are limited by suboptimal feature selection and pixel classification techniques.

Here we take a machine learning approach to the problem of exudates classification. In our experiments, the NN classifier is used as a baseline method for comparison with naive Bayes and SVM classifiers. Our results show that the naive Bayes and SVM classifiers perform substantially better than the NN classifier and the methods reported in our previous work. This is the first work to achieve practically useful exudate detection results on non-dilated fundus images.

Our methodology is described in section 2. In section 3, we report the results of each experiment. Finally, we discuss the results and conclude in section 4.

2. Methodology

We acquired 39 digital retinal images taken without pupil dilation from the Eye Care Center at Thammasat University Hospital. The images were captured in 24-bit color with a KOWA-7 non-mydratic retinal camera with a 45° field of view. We scaled all images to 752 × 500.

Firstly, we process the raw images using a pre-processing method described in section 2.1, in order to enhance the contrast of the image. We then eliminate the optic disc as described in section 2.2. Section 2.3 explains the features we extract from each pixel in a retinal image. Our feature selection and classification experiments with the naive Bayes and SVM classifiers are clarified in sections 2.4 and 2.5. In section 2.6, we explain how we obtain a baseline for comparison using a NN classifier. Finally, we describe the methods used to measure the system's performance in section 2.7.

2.1 Preprocessing

To obtain an image suitable for feature extraction, we perform several steps of preprocessing. First, we convert each original RGB image to grayscale by taking the average of the red, green, and blue channels. Second, we perform median filtering to remove shot noise. Finally, we apply "contrast-limited adaptive histogram equalization" (16) to enhance local contrast. Exudate and optic disc regions are typically much higher in intensity than neighboring image regions (17, 18), so the contrast enhancement technique tends to assign them the highest intensity values. The steps are illustrated in Figure 1.

2.2 Optic disc detection

After preprocessing, the optic disc has some characteristics similar to hard exudates: bright intensities and sharp boundaries. To prevent the optic disc from interfering with exudate detection, we first detect the optic disc and eliminate it from consideration.

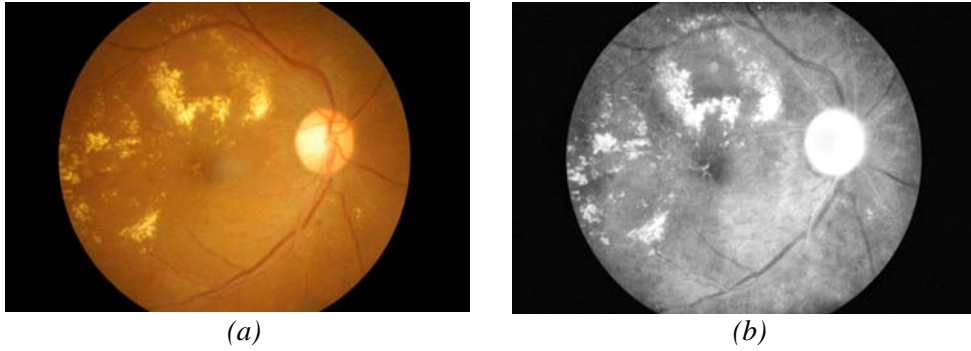


Figure 1. Image preprocessing. (a) Original retinal image. (b) The same image preprocessed by grayscale conversion, median filtering, and contrast enhancement.

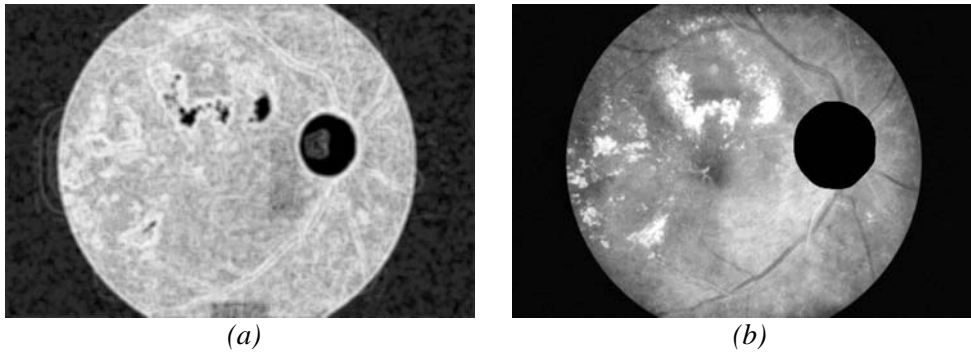


Figure 2. Optic disc detection result. (a) Entropy filtered image. (b) Optic disc area are eliminated from the contrast enhanced image.

We find that the optic disc is distinguished from the rest of the retina by its smooth texture. To determine which regions of the image are smooth or textured, for each point x , we obtain a probability mass function $P_{I_x} : \{0, 1, \dots, 255\} \mapsto [0..1]$ for the intensities I_x of the pixels in a local region around x then compute the entropy of that mass function.

$$H(I_x) = -\sum_{i=0}^{255} P_{I_x}(i) \cdot \log P_{I_x}(i) \quad (1)$$

This local pixel intensity entropy measure is high when the region around a pixel is complex and low when it is smooth. After filtering with the entropy operator we apply Otsu's binarization algorithm (19) to separate the complex regions from the smooth regions.

However, after binarization more than one candidate region is typically identified. We select the largest connected component R_i whose shape is approximately circular. We measure the circularity of R_i using the *compactness* measure

$$C(R_i) = 4\pi A(R_i) / P^2(R_i), \quad (2)$$

where $A(R_i)$ is the number of pixels in region i and $P(R_i)$ is the length of the boundary of R_i . To ensure that all pixels on the boundary of the optic disc are also grouped with the optic disc region, we apply binary dilation to the detected region. The results from this step are shown in Figure 2. This method for finding the optic disc achieves 100% accuracy on our data set. Eliminating the optic disc during preprocessing dramatically simplifies the decision that the classifier is required to make.

2.3 Feature extraction

Proteins and lipids leaking from damaged blood vessels form exudates. Exudates can be identified on an ophthalmoscope as areas with hard white or yellowish colors. They have varying sizes, shapes, and locations near the leaking capillaries within the retina.

We asked ophthalmologists how they identify exudates in an image so that our feature extraction would reflect ophthalmologists' expertise. We found that color, size, shape and texture are the most important features they consider. To differentiate exudate pixels from non-exudate pixels, we attempt to mimic ophthalmologists' expertise by extracting these relevant and significant features. As an initial set of candidate per-pixel features, we selected 15 features and used them as input for our classifier. A general motivation and an explanation of the motivation for each feature is explained in this section.

- (1) The pixel's *intensity* value after preprocessing. Exudate pixels can usually be distinguished from normal pixels by their intensity.
- (2) The *standard deviation* of the preprocessed intensity values in a window around the pixel. We use a window size of 15×15 . We use the standard deviation because exudates tend to be more highly textured than non-exudate regions, and the standard deviation is a simple indication of texture.
- (3) The pixel's *hue*. Hue characterizes chrominance or color information which should distinguish exudates from non-exudates.
- (4) The *number of edge pixels* in a region around the pixel. Exudates often form small clusters so they tend to have many edge pixels. We apply a Sobel edge operator then eliminate the strong edges arising

from blood vessels and the optic disc using decorrelation stretch (20) on the red band. We use a 17×17 neighborhood.

- (5) The *average intensity of the pixel's cluster*. We cluster the preprocessed image into contiguous regions with similar intensity using agglomerative clustering (21). This feature is simply the average of the intensities of the pixels assigned to the same cluster.
- (6) The *size of the pixel's cluster*, measured in pixels.
- (7) The *average intensity of the pixels in the neighborhood of the pixel's cluster*. The pixel's cluster is extracted and dilated using a 3×3 structuring matrix. The neighborhood pixels are those obtained by subtracting the original region from the dilated region.
- (8) The *ratio between the size of the pixel's cluster and the size of the optic disc*.
- (9) The *distance between the pixel's cluster and the optic disc*. We use the Euclidean distance between the centroid of the pixel's cluster and the centroid of the optic disc.
- (10) *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 (22). 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 $\sigma = 0.5$ from scale $\sigma = 1$, scale $\sigma = 1$ from $\sigma = 2$, scale $\sigma = 2$ from $\sigma = 4$, scale $\sigma = 4$ from $\sigma = 8$, scale $\sigma = 8$ from $\sigma = 16$,

and scale $\sigma = 16$ from $\sigma = 32$, respectively.

Before feature selection or classification, we z -scale (transform to a mean of 0 and a standard deviation of 1) all 15 features using the statistics of each feature over the training set. Examples of some of the features are shown in Figure 3.

2.4 Feature selection and classification using naive Bayes

The naive Bayes classifier uses the principle of Bayesian maximum a posteriori (MAP) classification: measure a finite set of features x_1, \dots, x_n then select the class

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

where

$$P(y | \mathbf{x}) \propto P(\mathbf{x} | y)P(y). \quad (3)$$

$P(\mathbf{x} | y)$ 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.

After z -scaling, all of our features x_i are continuous, but the simple version of naive Bayes just described requires discrete features, so we perform unsupervised proportional k -interval discretization as implemented in Weka (23). The technique uses equal-frequency binning, where the

number of bins is the square root of the number of values. After discretization, the conditional probability tables are obtained by simply counting over the training set.

Feature selection proceeds as follows. We first estimate the model of equation 3 from a training set using all 15 features, then we evaluate the resulting classifier's performance on a separate test set. We find that sensitivity is normally quite high with this approach, so we seek to reduce the number of false positives on the test set as much as possible by iteratively deleting features until the average of the precision and sensitivity ("PR," see section 2.7) 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 the classifier's PR improves, we permanently delete that feature then repeat the process. Finally, the best feature set and classifier are retained.

2.5 Feature selection and classification using support vector machines

Support vector machines 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:

$$\hat{y} = h(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^n \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b\right), \quad (4)$$

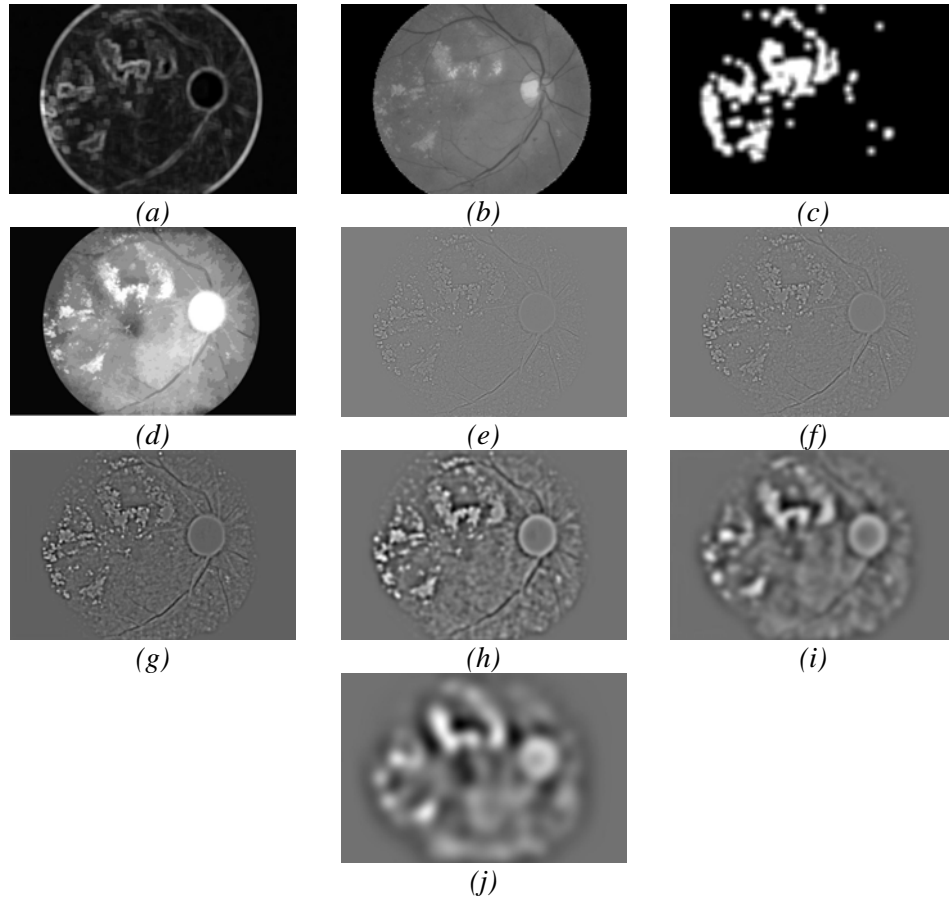


Figure 3. Input features. (a) Standard deviation of intensity. (b) Hue. (c) Number of edge pixels. (d) Cluster intensity. (e) DoG1. (f) DoG2. (g) DoG3. (h) DoG4. (i) DoG5. (j) DoG6.

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*.

In practice, the classes overlap in feature space and there is no perfectly separating hyperplane. In these cases a *soft margin* classifier is required. We use the ν -SVM (24) in which the parameter $\nu \in [0..1]$ controls how many support vectors are allowed to lie on the wrong side of the separating hyperplane.

We use the ν -SVM with a radial basis function (RBF) kernel of the form

$$K(x, x') = \exp(-\gamma \|x - x'\|). \quad (5)$$

Each support vector thus becomes the center of a RBF, and γ determines the area of influence that support vector has over the data space.

We use the best feature set obtained from naive Bayes as an initial feature set for the SVM. We then add features to the SVM 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.

This procedure is an adaptation of the standard greedy backward elimination feature selection method. Since training our SVM requires a two-dimensional search

over the hyperparameter space, full backward elimination is not practical. We thus use the naive Bayes model to select a candidate feature for elimination from the SVM. This means we only have to perform $O(N)$ hyperparameter searches, where N is the original number of features, rather than $O(N^2)$

2.6 Nearest neighbor classifier

As a simple baseline for comparison, we use a nearest neighbor classifier. The nearest neighbor classifier simply classifies a test instance with the class of the nearest training instance according to some distance measure. We experiment with two distance measures, Mahalanobis and Euclidean.

2.7 Performance measurement

We evaluate performance on the test set quantitatively by comparing the classifier's result to ground truth. To obtain ground truth, for each image, we used image processing software to hand label candidate exudate regions, then we asked an ophthalmologist to verify or reject each candidate region.

To evaluate classifier performance, we use sensitivity, specificity, precision, PR and accuracy on a per-pixel basis. All measures can be calculated based on four values, namely the true positive (TP) rate (the number of exudate pixels correctly detected), the false positive (FP) rate (the number of non-exudate pixels wrongly detected as exudate pixels), the false negative (FN) rate (the number of exudate pixels not detected), and the true negative (TN) rate (the number of non-exudate pixels correctly identified as non-exudate pixels). These values are defined in Table 1.

Table 1. Pixel-based evaluation.

	Disease status	
Test result	Present	Absent
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

From these quantities, the sensitivity, specificity, precision, PR and accuracy are computed using equations 6, 7, 8, 9, and 10 respectively. Sensitivity is the percentage of the actual exudate pixels that are detected, and specificity is the percentage of non-exudate pixels that are correctly classified as non-exudate pixels. Precision is the percentage of detected pixels that are actually exudate, and PR is the average of the precision and sensitivity. Accuracy is the overall per-pixel success rate of the classifier.

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

$$Specificity = \frac{TN}{TN + FP} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$PR = \frac{Precision + Sensitivity}{2} \quad (9)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (10)$$

3. Results

We performed preprocessing as described in section 2.1 on all 39 training and test images using Matlab. The window size for median filtering was 3×3 . For optic disc detection (section 2.2), we used a 9×9 neighborhood for entropy filtering, and an 11-pixel radius flat disc structuring element

for optic disc dilation. With a compactness threshold of 0.5, we obtained 100% accuracy in optic disc removal over all 39 images in our data set.

In our experiment, the classification is pixel-based, so each training or test sample represents one pixel in a training or testing image. For each image in the training set, we computed the features for every exudate pixel then randomly selected and computed features from an equal number of non-exudate pixels. The two sets of examples formed our training set. We thus obtained 231,734 samples, 115,867 examples of exudate pixels and 115,867 examples of non-exudate pixels, for training. From our test images, we used the 42,909 exudate pixels and 2,374,201 non-exudate pixels as a testing set.

We used Weka data mining software (23) running on a standard PC for feature discretization and naive Bayes classification. We used libSVM's (25) 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 was maximized.

We found that z -scaling normalization of the input features was necessary for Weka but slightly decreased SVM performance, so we report results with normalized input features for the naive Bayes classifier and unnormalized input features for the SVM classifier. The initially selected features, consisting of pixel's intensity value after preprocessing, standard deviation of the preprocessed intensity, pixel hue, number of edge pixels in a region around the pixel, average intensity of the pixel's cluster, size of the pixel's cluster, average intensity of the pixel in the neighborhood of the pixel's

cluster, ratio between the size of the pixel's cluster and the size of the optic disc, distance between the pixel's cluster and the optic disc, and the six difference of Gaussian filter responses, were fed into the naive Bayes classifier feature selection procedure (section 2.4). We then fed the best feature set from the naive Bayes classifier to the SVM classifier as an initial feature set and performed feature selection as described in section 2.5. Finally, we fed the best feature sets from the naive Bayes and SVM classifiers to the NN classifier (section 2.6) to obtain a performance baseline. For the NN classifier with Mahalanobis distance, we calculated the covariance matrix on the positive training examples only.

3.1 Experiment 1: naive Bayes classification

In our first experiment, we fit the naive Bayes model (equation 3) to the training set using all 15 features. The resulting classifier had an overall per-pixel sensitivity, specificity, precision, PR and accuracy of 95.84%, 96.56%, 33.49%, 64.67% and 96.55%, respectively. In the next step, we removed features from the classifier one by one and compared the resulting PR to PR obtained on the previous feature set. We obtained the best PR value (65.78%) by deleting cluster intensity, presumably due to its redundancy with the pixel intensity feature. We continued this process until the PR stopped improving. Finally, the best-performing classifier contained six features: the pixel's intensity after preprocessing, the standard deviation of the preprocessed intensities in a window around the pixel, the pixel hue, the number of edge pixels in a window around the pixel, the ratio between the size of the pixel's intensity cluster and the optic disc, and DoG4. The complete test results are listed in Table 2.

3.2 Experiment 2: SVM classification

In our second experiment, we performed a grid search to determine the best combination of hyperparameters ν (tolerance for training errors) and γ (radial basis function width) for the ν -SVM using the best feature set for the naive Bayes classifier as an initial feature set. We then added features back into the classifier one by one and repeated the grid search for each feature set combination. We found that PR fluctuated as we performed feature inclusion, so we continued including features until all 15 features were included. The results of the tests are shown in Table 3. The best performance was obtained using 10 features: the pixel's intensity after preprocessing, the standard deviation of the preprocessed intensities in a window around the pixel, the pixel hue, the number of edge pixels in a window around the pixel, the ratio between the size of the pixel's intensity cluster and the optic disc, distance between the pixel's cluster and the optic disc, DoG1, DoG2, DoG4, and DoG6, with $\gamma = 0.002$ and $\nu = 0.98$. This classifier has a sensitivity of 92.28%, specificity of 98.52%, precision of 53.05%, and PR of 72.67%. Its overall accuracy is 98.41%. Examples of exudate regions predicted by the SVM for two test set images are shown in Figure 4. Comparisons between the predicted exudates regions and the corresponding ground truth data are shown in Figure 5.

3.3 Experiment 3: NN classification

In a final experiment, we obtained a baseline for comparison using a NN classifier with Euclidean and Mahalanobis distance metrics. To compare with the performance of the the naive Bayes and

SVM classifiers, we used the best feature sets obtained for naive Bayes and the SVM. On best feature set obtained from the naive Bayes classifier, the NN classifier with Euclidean and Mahalanobis distance have a PR of 61.54% and 61.81%, respectively. On the best feature set obtained from the SVM classifier, the NN classifier with Euclidean and Mahalanobis distance obtained a PR of 65.15% and 64.99%, respectively. The results are compared in Table 4. The results indicate that the naive Bayes and SVM classifiers perform substantially better in PR than the NN classifier. In addition, the NN classifier using the best feature set obtained from the SVM classifier performs better than that using the best feature set for the naive Bayes classifier.

4. Discussion and conclusion

In this paper, we propose 15 per-pixel features as potential indicators of exudates. We use naive Bayes and SVMs for feature selection and pixel classification. We filter images for noise, enhance image contrast, detect and remove the optic disc, extract local features describing pixels or regions, then classify those features using a model built from a training set. The naive Bayes classifier, after feature selection, achieves an overall per-pixel sensitivity of 93.38%, specificity of 98.14%, precision of 47.51%, PR of 70.45% and an overall accuracy of 98.05% on a test set not used during training. The SVMs classifier achieves an overall per-pixel sensitivity of 92.28%, specificity of 98.52%, precision of 53.05%, PR of 76.27% and an overall accuracy of 98.41%.

Table 2. Naive Bayes prediction performance

Features	Sensitivity	Specificity	Precision	PR	Accuracy
All features	95.84%	96.56%	33.49%	64.67%	96.55%
Without CI	95.68%	96.91%	35.89%	65.78%	96.89%
Without CI and DoG 3	95.58%	97.12%	37.51%	66.55%	97.09%
Without CI, DoG 3 and NI	95.20%	97.36%	39.49%	67.35%	97.33%
Without CI, DoG 3, NI and CS	96.10%	97.41%	40.17%	68.14%	97.39%
Without CI, DoG 3, NI, CS and DoG 5	95.76%	97.62%	42.12%	68.94%	97.59%
Without CI, DoG 3, NI, CS, DoG 5 and DoG 2	95.41%	97.80%	43.98%	69.70%	97.76%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2 and distance	94.82%	97.95%	45.56%	70.19%	97.90%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance and DoG 6	93.82%	98.06%	46.62%	70.22%	97.98%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6 and DoG 1	93.38%	98.14%	47.51%	70.45%	98.05%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and NE	92.12%	96.46%	31.98%	62.05%	96.38%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and SD	90.03%	97.70%	41.48%	65.76%	97.57%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and SR	93.70%	97.52%	40.61%	67.16%	97.46%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and intensity	92.08%	97.81%	43.18%	67.63%	97.71%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and hue	94.14%	97.77%	43.27%	68.70%	97.70%
Without CI, DoG 3, NI, CS, DoG 5, DoG 2, distance, DoG 6, DoG 1 and DoG 4	91.42%	98.11%	46.60%	69.01%	97.99%

* CI = cluster intensity, CS = cluster size, DoG = difference of Gaussian, NE = number of edge pixels, NI = neighborhood intensity, SD = standard deviation and SR = size ratio

Table 3. SVM prediction performance

Features	Sensitivity	Specificity	Precision	PR	Accuracy
*Best feature set from NB ($\gamma=0.004$, $\nu=0.999$)	87.90%	98.57%	52.56%	70.22%	98.37%
With DoG 1 ($\gamma=0.004$, $\nu=0.999$)	88.71%	98.60%	53.38%	71.05%	98.42%
With DoG 1 and DoG 6 ($\gamma=0.004$, $\nu=0.99$)	92.54%	98.21%	48.24%	70.38%	98.10%
With DoG 1, DoG 6 and distance ($\gamma=0.004$, $\nu=0.995$)	93.00%	98.32%	50.06%	71.53%	98.23%
With DoG 1, DoG 6, distance and DoG 2 ($\gamma=0.002$, $\nu=0.98$)	92.28%	98.52%	53.05%	72.67%	98.41%
With DoG 1, DoG 6, distance, DoG 2 and DoG 5 ($\gamma=0.002$, $\nu=0.98$)	92.89%	98.43%	51.61%	72.25%	98.33%
With DoG 1, DoG 6, distance, DoG 2, DoG 5 and CS ($\gamma=0.004$, $\nu=0.995$)	93.61%	98.17%	47.98%	70.79%	98.08%
With DoG 1, DoG 6, distance, DoG 2, DoG 5, CS and NI ($\gamma=0.004$, $\nu=0.9$)	93.35%	97.89%	44.38%	68.87%	97.81%
With DoG 1, DoG 6, distance, DoG 2, DoG 5, CS, NI and DoG 3 ($\gamma=0.004$, $\nu=0.995$)	92.36%	98.15%	47.44%	69.90%	98.05%
With DoG 1, DoG 6, distance, DoG 2, DoG 5, CS, NI, DoG 3 and CI ($\gamma=0.004$, $\nu=0.995$)	92.18%	97.64%	41.36%	66.67%	97.54%

* Best feature set from NB is pixel's intensity, standard deviation, pixel's hue, number of edge pixels, ratio between the size of the pixel's cluster and the size of the optic disc and DoG 4

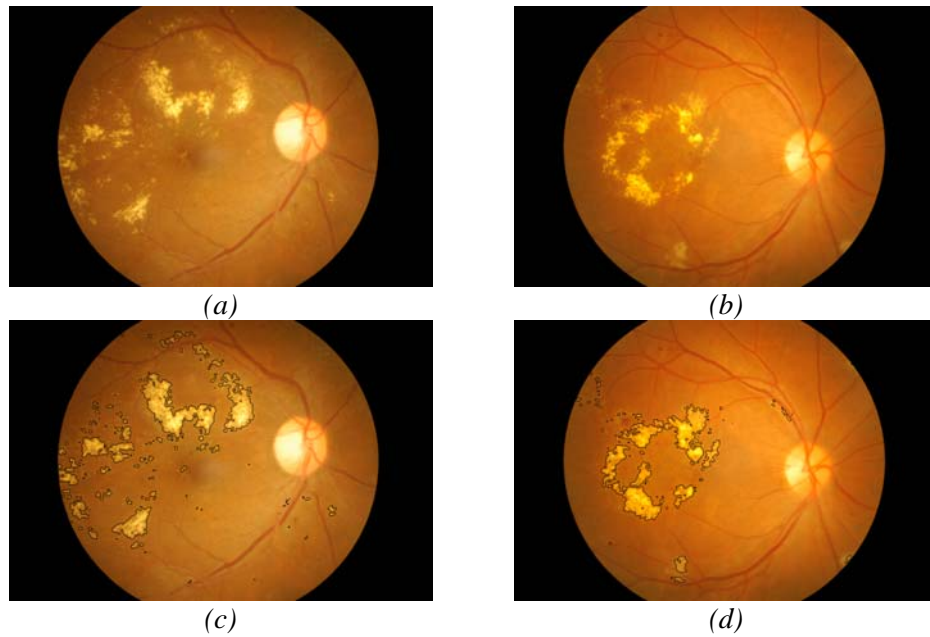


Figure 4. Exudates detection by the SVM classifier. (a–b) Two original test set images. (c–d) Perimeters of the exudate regions predicted by the SVM, superimposed on images (a) and (b).

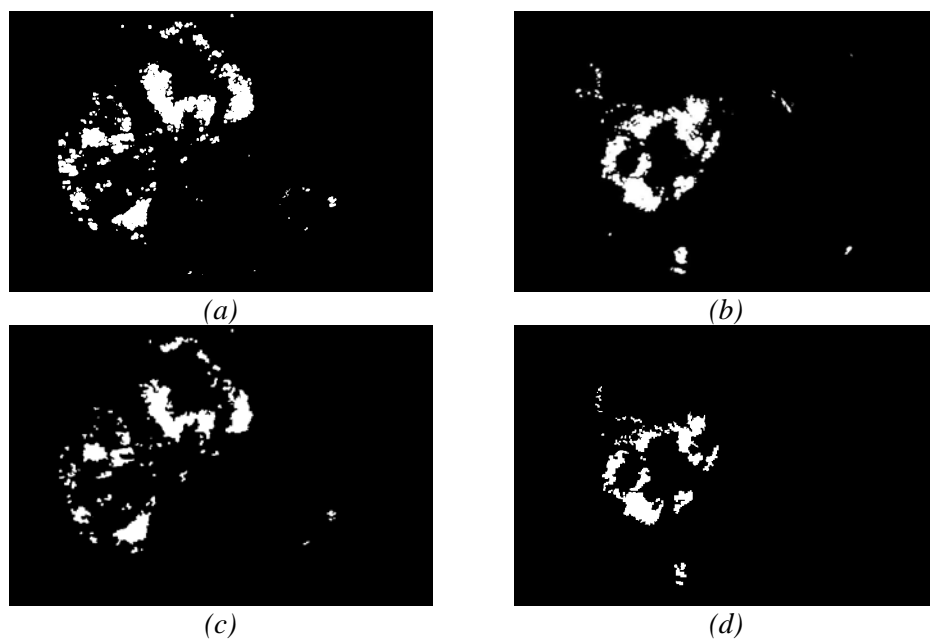


Figure 5. Comparison of SVM prediction and ground truth. (a–b) Exudate pixels predicted by the SVM for the original images from Figures 4a and 4b. (c–d) Ground truth data for the same two images.

Table 4. PR performance comparison

Classifier	NB	SVM	NN (Euc.)	NN (Mah.)
Best feature set for NB	70.45%	70.22%	61.54%	61.81%
Best feature set for SVM	68.94%	72.67%	65.15%	64.99%

The final SVM classifier contains ten features: the pixel's intensity after preprocessing, the standard deviation of the preprocessed intensities in a window around the pixel, the pixel hue, the number of edge pixels in a window around the pixel, the ratio between the size of the pixel's intensity cluster and the optic disc, the distance between the pixel's cluster and the optic disc and the response at the pixel to a derivative of Gaussian filter. Both the naive Bayes and SVM classifiers perform well compared to a baseline nearest neighbor classifier.

The naive Bayes classifier benefits from aggressive feature selection, probably because it assumes conditional independence of the features even though some of our features are clearly highly correlated. The support vector machine classifier is better able to handle these statistical dependencies, achieving slightly higher performance with a larger feature set at the cost of substantially increased compute time during both training and testing. It might be possible to achieve similar performance with faster classifiers such as Adaboost cascades.

How do the exudates predicted by naive Bayes and the SVM compare? Figure 6 shows closeups of the classifiers' exudate boundary predictions for two typical exudates. The SVM classifier tends to delineate exudate boundaries more accurately with fewer false detections. Although the SVM is only slightly better than naive Bayes in terms of the quantitative performance evaluation, its superiority is very clear in the images.

This shows that two methods with quantitatively similar per-pixel accuracy can perform differently in terms of qualitative accuracy. The

reason is that false detections near true exudate boundaries are qualitatively less serious than isolated false detections. Similarly, a group of false negatives along the boundary of a true exudate is qualitatively less serious than a completely missed exudate region. In principle, we could define a new performance measure that weighs false detections and false negatives differently depending on their distance from or connectedness to true exudate regions. This could quantify the performance of the SVM-based classifier. However, here we simply use the standard performance measures and show the qualitative difference in the images.

Both the naive Bayes classifier and the SVM occasionally miss faint exudates and incorrectly detect as exudate image artifacts or retinal structures that share some characteristics with exudates. For example, as shown in Figure 7, strong, high contrast choroidal blood vessels appearing in the retinal background can be incorrectly detected as exudates. As another example, Figure 8 shows that faint blood vessels are also sometimes incorrectly detected as exudates. Overall, our experimental results show that careful preprocessing, good features, feature selection, and an appropriate classifier together provide excellent exudates detection performance even on retinal images acquired from patients whose pupils have not been dilated. We use a relatively large set of pixel descriptor features along with a principled method for feature selection, and we compare three different machine learning techniques with different characteristics. To our knowledge, the results reported here are the best automatic exudate detection results thus far. Additionally, our data set is apparently the largest existing set with hand-labeled

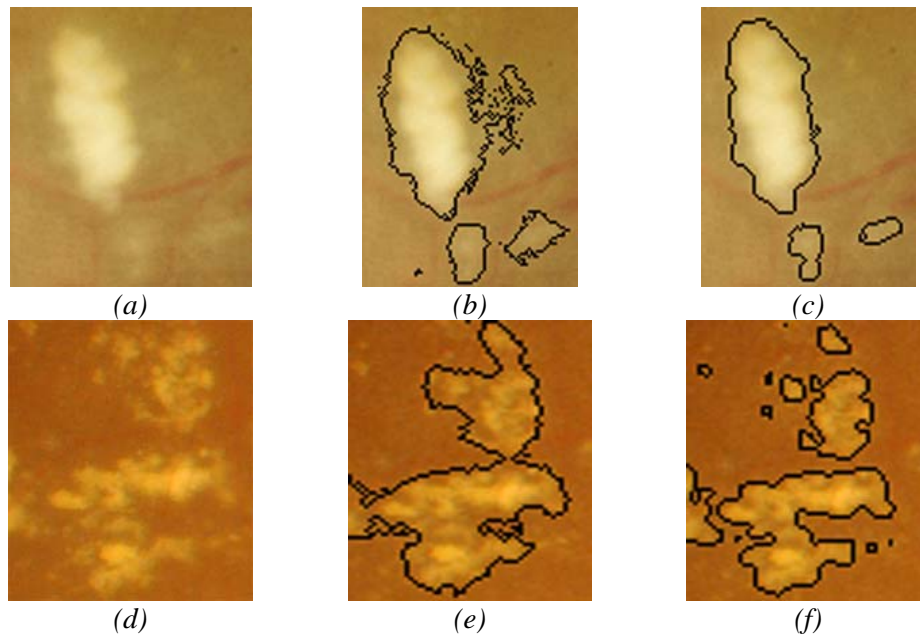


Figure 6. Closeups of results for two typical exudates. (a) and (d) Original images. (b) and (e) Naive Bayes classification results. (c) and (f) SVM classification results.

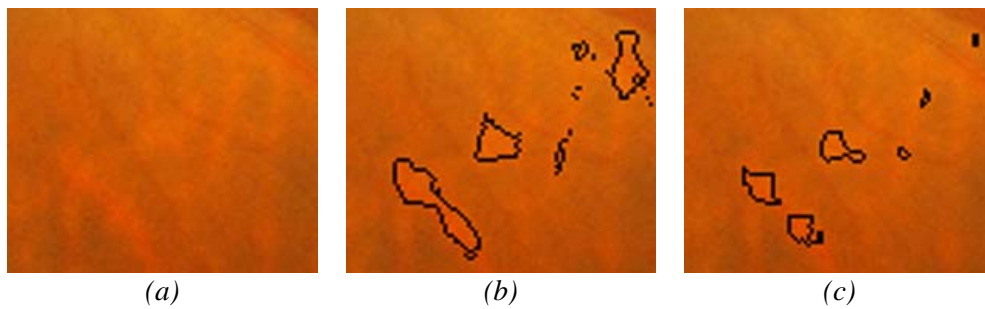


Figure 7. Example false detection of exudates on choroidal blood vessel. (a) Original image. (b) Detection results for naive Bayes classifier. (c) Detection results for SVM classifier.

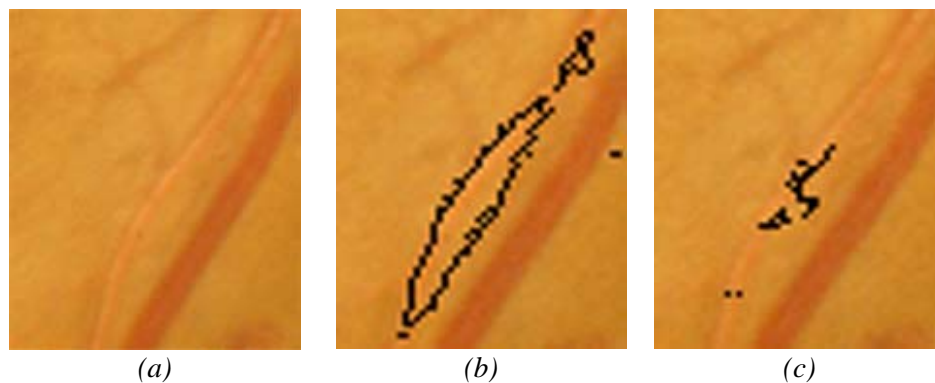


Figure 8. Example false detection of exudates on faint blood vessel. (a) Original image. (b) Detection results for naive Bayes classifier. (c) Detection results for SVM classifier.

ground truth provided by doctors. The data set and SVM classifier (in the standard libSVM format) are available to interested researchers at <http://ict.siit.tu.ac.th/~ivc>.

Although care must be taken not to overgeneralize the applicability of these results, as the feature set and SVM hyperparameters ν and γ have been tuned to maximize performance on the test set, and validation on a new independent test set is still necessary, this work is a promising step towards automated diagnosis of diabetic patients' retinal images. Automated detection of exudates in diabetic patients' retinas could help enable early detection of diabetic retinopathy and could help doctors track the progress of treatment over time.

In future work we plan to expand the data set and explore using the system as a practical aid to help ophthalmologists screen patients for diabetic retinopathy symptoms quickly and easily.

Acknowledgments

We thank Thammasat University Hospital for the images used in these experiments, and we thank Thammasat University medical doctors for providing ground truth data. This research was supported by the Thailand National Electronics and Computer Technology Center (NECTEC) and by Thailand Research Fund grant MRG4780209 to MND.

References

- [1] Liu, Z.; Chutatape, O.; Krishna, S.M. Presented at the 19th IEEE Conference on Engineering in Medicine and Biology Society, Chicago, USA, Oct 30-Nov 2, 1997.
- [2] Ege, B.; Hejlesen, O.; Larsen, O.; Moller, K.; Jennings, B.; Kerr, D.; Cavan, D.A. *Comput. Meth. Programs Biomed.* **2000**, *62*, 165–175.
- [3] Sinthanayothin, C.; Boyce, J.F.; Williamson, T.H.; Cook, H.L.; Mensah, E.; Lal, S.; Usher, D. *J. Diabet. Med.* **2002**, *19*, 105–112.
- [4] Usher, D.; Dumskyj, M.; Himaga, M.; Williamson, T.H.; Nussey, S.; Boyce, J. *J. Diabet. Med.* **2004**, *21*, 84–90.
- [5] Kavitha, D.; Shenbaga, S.D. Presented at the 2nd ICISIP Conference on Intelligent Sensing and Information Processing, Madras, India, Jan 4-7, 2005.
- [6] Osareh, A.; Mirmehdi, M.; Thomas, B.; Markham, R. In *Medical Image Understanding and Analysis Conference*, Claridge, E., Bamber, J. Eds.; BMVC Press: UK, 2001; pp 49–52.
- [7] Osareh, A.; Mirmehdi, M.; Thomas, B.T.; Markham, R. In: *MICCAI '02: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II*, Springer-Verlag: London, UK, 2002; pp 413–420.
- [8] Zhang, X.; Chutatape, O. Presented at the IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, June 20-25, 2005.
- [9] Katarzyna, S.; Adam, S.; Radim, C.; Georg, M. Segmentation of Fundus Eye Images Using Methods of Mathematical Morphology for Glaucoma Diagnosis, Lecture Notes in Computer Science, 2004.
- [10] Sanchez, C.I.; Hornero, R.; Lopez, M.I.; Poza, J. Presented at the 26th IEEE Conference on Engineering in Medicine and Biology Society, California, USA, Sep 1-4, 2004.
- [11] Walter, T.; Klevin, J.C.; Massin, P.; Erginay, A. *IEEE Trans. Med. Imaging* **2002**, *21*, 1236–1243.
- [12] Wang, H.; Hsu, W.; Goh, K.G.; Lee, M.L. Presented at IEEE Conference on Computer Vision and Pattern Recognition, South Carolina, USA, June 13-15, 2000.
- [13] Sopharak, A.; Uyyanonvara, B. Presented at the Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information, Ubon-ratchathani, Thailand, May 10-13, 2006.

- [14] Sopharak, A.; Uyyanonvara, B. Presented at the Third Conference on World Congress on Bioengineering, Bangkok, Thailand, July 19-11, 2007.
- [15] Sopharak, A.; Uyyanonvara, B. Presented at the Conference on Advances in Computer Science and Technology, Phuket, Thailand, April 2-4, 2007.
- [16] Gonzalez, R.C.; Woods, R.E.; Eddins, S.L. *Digital Image Processing Using Matlab*. Prentice Hall: Upper Saddle River, NJ, 2003.
- [17] Osareh, A.; Mirmehdi, M.; Thomas, B.; Markham, R. *Br. J. Ophthalmol* **2003**, *87*, 1220–1223.
- [18] Sinthanayothin, C.; Boyce, J.F.; Cook, H.L.; Williamson, T.H. *Br. J. Ophthalmol* **1999**, *83*, 231–238.
- [19] Otsu, N. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66.
- [20] Phung, S.L.; Bouzerdoum, A.; Chai, D. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 148–154.
- [21] Forsyth, D.A.; Ponce, J. *Computer Vision: A Modern Approach*; Prentice Hall: New Jersey, 2003; pp 433-462.
- [22] Kenneth, R.S.; John, C.R.; Matthew, J.P.; Thomas, J.F.; Michael W.D. 2006.<http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/diffgaussians/index.html> (accessed July 30, 2007).
- [23] Witten, I.H.; Frank, E. *Data Mining: Practical machine learning tools and techniques*, 2nd ed.; San Francisco: Morgan Kaufmann, 2005.
- [24] Scholkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. *Neural Comput.* **2000**, *12*, 1207-1245.
- [25] Chang, C.C.; Lin, C.J. *LIBSVM: A Library For Support Vector Machines*. 2001.<http://www.csie.ntu.edu.tw/~cjlin/libsvm> (accessed Aug 1, 2007).