Classification of Capsule Endoscope Images Using Local Binary Patterns and Support Vector Machines

Edreen Bryan Valdeavilla[1], Shaou-Gang Miaou[2]

Department of Electronic Engineering, Chung Yuan Christian University
200 Chung Pei Rd. Chungli, Taiwan 32023, R.O.C.
[1]edreenbryan@gmail.com
[2]sgmiaou@gmail.com

***Abstract*—Wireless capsule endoscopy (WCE) is the state-of-the-art technology used for the investigation of intestines. With this technology, a physician can examine the entire section of the intestines, including the blind section not reachable with a traditional endoscope. However, one problem derived from this new technology is the tremendous amount of images that need to be inspected by human eyes and this becomes a burden to the physician. This paper presents some methods for an automatic detection system to identify suspected capsule endoscope images such as chyme blocked, suspected blood indicator, and ulcer in order to reduce this burden. These methods use standard local binary pattern (LBP) and Sobel-LBP as recognition features and support vector machines (SVM), nu-SVM (*v*-SVM), and one class SVM (OCSVM) as classifiers. For comparison, the nearest neighbor classifier is also considered. Experimental results carried out on 10-runs of 5-fold cross validation using 80% of the training dataset showed that the combination of image downsampling by 1, standard LBP, and the OCSVM classifier outperforms other methods. The resulting accuracy is 98.10% and the computational time in the testing phase is below 0.5 seconds per image, which is applicable for practical applications.**

***Keywords*—Support Vector Machines, Wireless Capsule Endoscopy, Local Binary Pattern, Cross Validation, Arithmetic Means.**

1. INTRODUCTION

Wireless capsule endoscopy (WCE) is a disposable imaging capsule which could visualize the small bowel mucosa and trace the abnormalities of the small bowel. This capsule is able to transmit color and high fidelity images of gastrointestinal tract, including intestines [1].

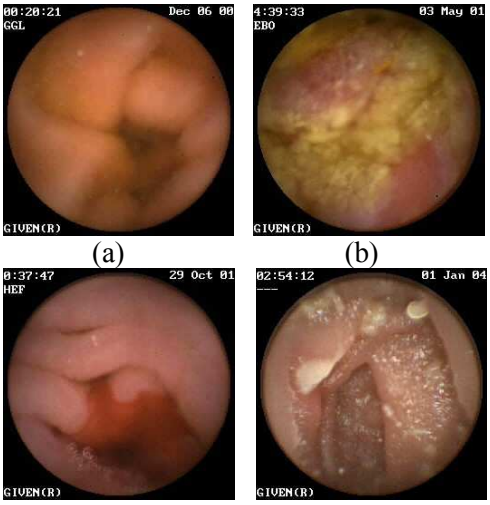
An image sequence with two frames per second and 256 x 256 resolution is transmitted by a typical capsule endoscope for about 7 to 8 hours, generating a total of approximately 50,000 frames for each examination. The image sequence is then viewed and annotated by a physician [1], [2]. The physician normally takes 45 minutes to 2 hours to view and analyse the result of one examination, depending on the physician's experience and the patient's abnormalities.

This paper proposes an automatic detection system to identify the image with either chyme blocked, suspected blood indicator (SBI), or white spot abnormality as shown in Fig. 1. These three abnormalities are selected because they occur often in gastrointestinal tract.

The block diagram of the system is shown in Fig. 2. First, the endoscopy images are downsampled. Second, the region of interest (ROI) in downsampled images is extracted. Third, Image features are extracted using the average intensity, followed by the generation of local binary pattern (LBP) histogram. Two kinds of LBP are examined in this paper, which are the standard LBP and the Sobel-LBP. After that, a support vector machine (SVM) training or testing is conducted. Three kinds of SVM are examined in this paper, which are the basic SVM, *v*-SVM, and one-class SVM (OCSVM).

The LBP is a nonparametric local texture descriptor that can be derived efficiently [3]. In addition, the LBP texture features are invariant to any monotonic change in gray level intensities, resulting in more robust representation of textures under varying illumination conditions [4]. SVM is a supervised learning method for classification [5]. The learning strategy of SVM is based on the principle of structural risk minimization, which makes SVM has better generalization ability than other traditional learning machines that are based on the learning principle of empirical risk minimization [6]. However, due to the fact that the sizes of normal and abnormal images acquired are different, the

problem of class-boundary-skew (CBS) may occur when a basic SVM is used. To solve this problem, this paper also considers two new variants of SVM, which are called ν-SVM [7], [8] and one class SVM (OCSVM) [9], [10].



(a) (b)

(c) (d)

Fig. 1 Example of the capsule endoscopy images: (a) normal, (b) chyme blocked, (c) suspected blood indicator (SBI), and (d) white spots.

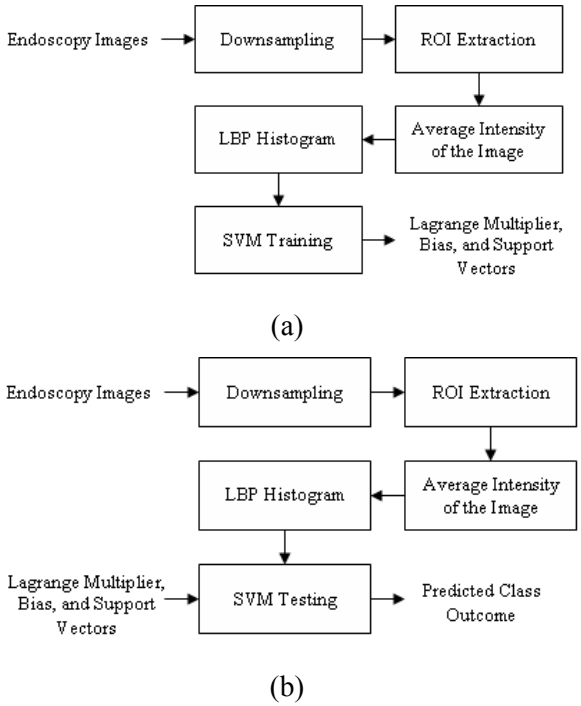


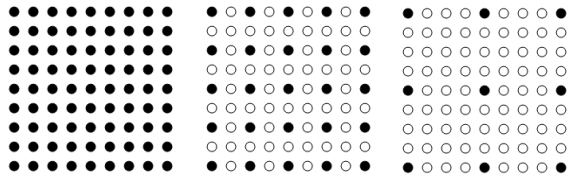
Fig. 2 System block diagram: (a) training process and (b) testing process.

This paper is organized as follows. Section 2 describes image downsampling and the extraction of region of interest (ROI). Section 3 describes the methods of standard LBP and Sobel-LBP. Section 4 introduces SVM classifiers, including basic SVM, *v*-SVM, and OCSVM. Section 5 presents the experimental results. Finally, conclusions are drawn in Section 6.

2. IMAGE DOWNSAMPLING AND REGION OF INTEREST (ROI)

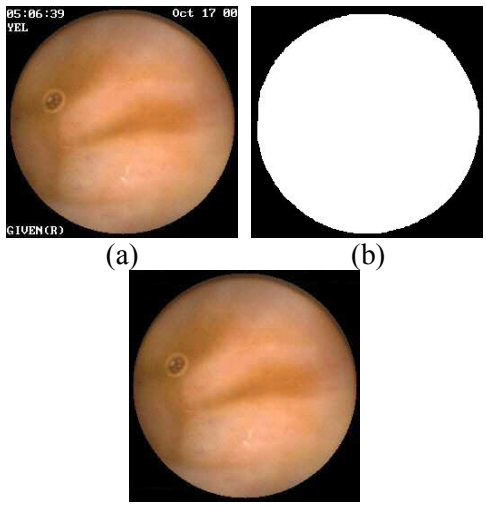
In this paper, the three downsampling factors considered are 1, 2 and 4 which are shown in Fig. 3. A downsampling factor of 1 means that the original image with 256 x 256 pixels would be used for the next block (i.e. feature extractor). The downsampling process is two-dimensional. Thus, only one quarter and one sixteenth of the original pixels will be preserved if downsampling factors of 2 and 4 are performed, respectively. Only the preserved pixels will be processed in the next block.

In addition, only the pixels inside an ROI are considered in every feature extractor. The ROI used here is a circular region located in the middle and obtained by excluding all the texts and the pure black background from a capsule endoscopy image as shown in Fig. 4.



(a) (b) (c)

Fig. 3 Image downsampling by a factor of (a) 1, (b) 2, and (c) 4. Here, only pixels denoted as black dots will be processed in the next step.



(a) (b)

(c)

Fig. 4 ROI extraction: (a) original image, (b) binarized image with a circular ROI, and (c) image after ROI extraction.

3. LOCAL BINARY PATTERNS (LBP)

LBP operators offer an alternative approach for texture representation. They are computationally efficient and nonparametric local image texture descriptors. In addition, LBP texture features are invariant to any monotonic change in gray level intensities.

3.1. Standard LBP

The standard LBP operator takes a local neighborhood around each pixel, thresholds the pixels of the neighbourhood at the value of the central pixel and uses the resulting binary-valued image patch as a local image descriptor [3]. It was originally defined for 3 x 3 neighborhoods, giving 8 bit codes based on the 8 pixels around the central one. Formally, the standard LBP operator takes the form

$$LBP = \sum_{n=0}^{7} 2^n s(i_n - i_c) \quad (1)$$

where in this case n runs over the 8 neighbors of the central pixel c, i_n and i_c are the intensity values at n and c, respectively, and

$$s(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

The LBP encoding process is illustrated in Fig. 5.

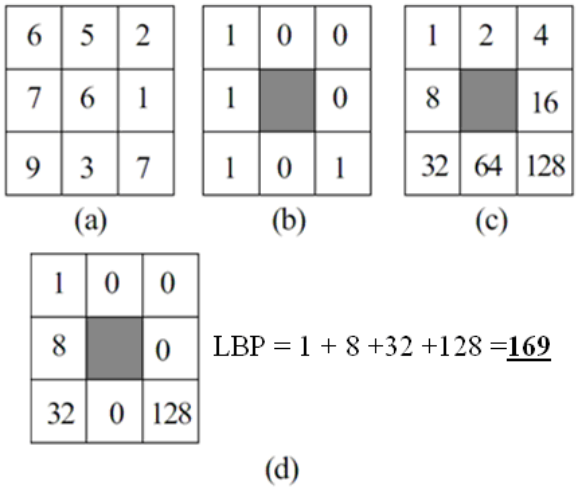


Fig. 5 An example of LBP value estimation: (a) an original 3 x 3 neighborhood, (b) the values of the pixels in the thresholded neighborhood, (c) binomial weights assigned to the corresponding pixels, and (d) the values of eight neighbor pixels are summed to obtain a single value for the corresponding pattern.

3.2. Sobel-LBP

The Sobel operator contains two 3x3 kernels (horizontal kernel S_x and vertical kernel S_y) which are convolved with the original image I to obtain gradient approximations [12]:

$$I^x = S_x * I = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * I, \qquad I^y = S_y * I = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I, \quad (3)$$

where I^x and I^y represent the horizontally and vertically filtered results, respectively. Normally I^x and I^y are combined to give the gradient magnitude $\sqrt{(I^x)^2 + (I^y)^2}$. Here the Sobel-LBP operator is defined as the concatenation of LBP operations on I^x and I^y:

$$Sobel - LBP = \{Sobel - LBP^x, Sobel - LBP^y\}, \quad (4)$$

where

$$Sobel - LBP^x = \sum_{n=0}^{7} 2^n s(I_n^x - I_c^x), \qquad Sobel - LBP^y = \sum_{n=0}^{7} 2^n s(I_n^y - I_c^y) \quad (5)$$

where in this case n runs over the 8 neighbors of the central pixel c, I_n^x and I_c^x are the intensity values of horizontally filtered results at n and c, I_n^y and I_c^y are the intensity values of vertically filtered results at n and c, respectively.

4. SUPPORT VECTOR MACHINES (SVM)

The main idea of a support vector machine is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative training samples is maximized [11], as shown in Fig. 6.

4.1. Support Vector Machine (SVM)

Let $\{(\mathbf{x}_i, y_i) \mid i = 1, 2, \ldots, N\}$ be the training sample, where $\mathbf{x}_i$ is the i-th input sample, y_i is the corresponding target output, and N is the number of training samples. We assume that the sample represented by the subset $y_i = +1$ and the sample represented by the subset $y_i = -1$ are "linearly separable." Then, the equation of a decision surface in the form of a hyperplane is given by [5], [11]:

$$\mathbf{w}^T\mathbf{x}+b=0 \tag{6}$$

where $\mathbf{w}$ is an adjustable weight vector, $\mathbf{x}$ is an input vector, and b is a bias. For linearly separable patterns, we obtain:

$$\mathbf{w}^T\mathbf{x}+b\geq 0 \quad \text{for } y_i=+1 \tag{7}$$
$$\mathbf{w}^T\mathbf{x}+b<0 \quad \text{for } y_i=-1$$

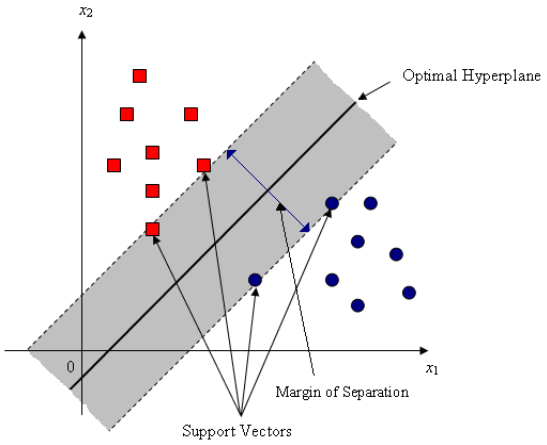


Fig. 6 Illustration of the concept of SVM.

When the data are not separable, a soft margin classifier is used. This introduces a misclassification cost C, which is assigned to each misclassified training sample. This can be done by introducing positive slack variables ξ_i as follows [5], [11]:

$$y_i(\mathbf{w}^T\mathbf{x}_i+b)\geq 1-\xi_i,\forall i$$
$$\xi_i\geq 0,\forall i \tag{8}$$

If a classification error occurs, the corresponding ξ_i must exceed unity, so $\sum_i\xi_i$ is an upper bound for the number of classification errors. Then, the cost function $\Phi(\cdot)$ to be minimized is given by:

$$\Phi(\mathbf{w},\xi_i)=\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{i=1}^{N}\xi_i \tag{9}$$

where C is a user-specified positive parameter. The parameter C controls the trade-off between complexity of the machine and the number of nonseparable points. A larger C corresponds to a higher penalty of classification errors.

By taking partial differential to the Lagrangian with respect to the variables, the dual problem becomes:

$$L_D(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k(\mathbf{x}_i,\mathbf{x}_j) \tag{10}$$

subject to

$$0\leq\alpha_i\leq C,\quad \sum_{i=1}^{N}\alpha_i y_i=0,\quad i=1,2,\ldots,N \tag{11}$$

In this paper, a radial basis function (RBF) is used as the kernel function. The RBF kernel is defined as

$$k(\mathbf{x}_i,\mathbf{x}_j)=\exp\left(-\frac{1}{2\sigma^2}\left\|\mathbf{x}_i-\mathbf{x}_j\right\|\right) \tag{12}$$

where σ is a user-specified parameter.

For an unseen data vector $\mathbf{x}$, the decision function is given by:

$$D(\mathbf{x})=sign\left(\sum_{i=1}^{N}\alpha_i y_i k(\mathbf{x}_i,\mathbf{x})+b\right) \tag{13}$$

4.2. *v*-SVM

The v-SVM uses a new parameter v which controls the number of support vectors and training errors [7], [8], [13]. The parameter $v\in(0, 1]$ is an upper bound of the fraction of training errors and a lower bound of the fraction of support vectors.

Given training vectors $\mathbf{x}_i\in R^n$, $i=1, 2, \ldots, N$, and the corresponding target output $\mathbf{y}\in R^N$ such that $y_i\in\{1, -1\}$, the primal problem becomes:

$$\Phi(\mathbf{w},\xi_i)=\frac{1}{2}\mathbf{w}^T\mathbf{w}-v\rho+\frac{1}{N}\sum_{i=1}^{N}\xi_i \tag{14}$$

subject to

$$y_i(\mathbf{w}^T\mathbf{x}_i+b)\geq\rho-\xi_i,$$
$$\xi_i\geq 0,\ i=1,2,\ldots,N,\quad \rho\geq 0. \tag{15}$$

By taking partial differential to the Lagrangian with respect to the variables, the dual problem becomes:

$$L_D(\alpha)=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k(\mathbf{x}_i,\mathbf{x}_j) \tag{16}$$

subject to

$$0\leq\alpha_i\leq 1,\quad \sum_{i=1}^{N}\alpha_i=vN,$$
$$\sum_{i=1}^{N}\alpha_i y_i=0,\quad i=1, 2, \ldots, N \tag{17}$$

Again, RBF is used as the kernel function. The RBF kernel is defined as

$$k(\mathbf{x}_i,\mathbf{x}_j)=\exp(-\gamma\,\|\,\mathbf{x}_i-\mathbf{x}_j\,\|) \tag{18}$$

where γ is the kernel width parameter.

For an unseen data vector $\mathbf{x}$, the decision function is given by:

$$D(\mathbf{x})=sign\left(\sum_{i=1}^{N}y_i\left(\frac{\alpha_i}{\rho}\right)(k(\mathbf{x}_i,\mathbf{x})+b)\right) \tag{19}$$

4.3. One Class SVM (OCSVM)

One-class SVM is used for estimating the support of a high-dimensional distribution [9], [13], [14]. It is an unsupervised approach that

separates outliers from the majority. This method seeks an approximation function to categorize the majority of data [19]. The OCSVM algorithm maps input data into a high dimensional feature space (via a kernel) and iteratively finds the maximal margin hyperplane which best separates the training data from the origin as shown in Fig. 7. The OCSVM may be viewed as a regular two-class SVM where all the training data lie in the first class, and the origin is taken as the only member of the second class [20].

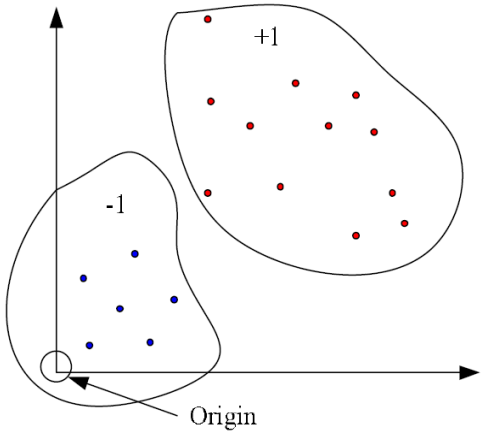


Fig. 7 Geometry interpretation of OCSVM [14].

Given training vectors $\mathbf{x}_i \in R^n$, $i = 1, 2, \ldots, N$ without any class information, the primal problem becomes:

$$\Phi(\mathbf{w}, \xi_i) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \rho + \frac{1}{\nu N}\sum_{i=1}^{N}\xi_i \qquad (20)$$

subject to

$$\mathbf{w}^T\mathbf{x}_i \geq \rho - \xi_i,$$
$$\xi_i \geq 0,\ i = 1, 2, \ldots, N. \qquad (21)$$

By taking partial differential to the Lagrangian with respect to the variables, the dual problem is similar to that in (17) subject to

$$0 \leq \alpha_i \leq 1,\ i = 1, 2, \ldots, N,\ \sum_{i=1}^{N}\alpha_i = \nu N. \qquad (22)$$

Again, RBF is used as the kernel function (Eq. (19)).

For an unseen data vector $\mathbf{x}$, the decision function is given by:

$$D(\mathbf{x}) = sign\left(\sum_{i=1}^{N}\alpha_i k(\mathbf{x}_i, \mathbf{x}) - \rho\right) \qquad (23)$$

5. EXPERIMENTAL RESULTS

The experiment used 350 endoscopy images, including 136 normal images and 214 abnormal images (150 images with chyme blocked and/or SBI, and the remaining 64 images with white spot lesions). 10-runs of cross-validation using 35 (10%), 44 (12.5%), 70 (20%), 175 (50%), and 280 (80%) training datasets, were conducted to determine the classifier parameters (which were determined from a wide range of numbers) and to evaluate the recognition accuracy using a-means (to be defined later). In this paper, the capsule endoscope images without abnormalities were considered as positive data.

The performance of a classifier can be determined using the quantities of data samples defined by the confusion matrix (Table 1).

TABLE 1
THE CONFUSION MATRIX

		Actual	
		Positive	**Negative**
Predicted	**Positive**	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

The entries in the confusion matrix in this paper have the following meaning:

- *TP* is the number of correct classification of normal images.
- *FP* is the number of incorrect classification of abnormal images.
- *TN* is the number of correct classification of abnormal images.
- *FN* is the number of incorrect classification of normal images.

The true positive rate (TP rate) or sensitivity or recall is the proportion of positive cases that were correctly identified and is given by:

$$\text{TP rate} = \text{sensitivity} = \text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (24)$$

The true negative rate (TP rate) or specificity is the proportion of negative cases that were correctly identified and is given by:

$$\text{TN rate} = \text{specificity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (25)$$

The recognition accuracy was evaluated by the arithmetic means (a-means) [18], defined as

$$a = \frac{1}{2}(\text{TP rate} + \text{TN rate}) \qquad (26)$$

The a-means was chosen as the evaluator since it is a reliable performance index, especially for the imbalanced test data set. For balanced test data set, this a-means is equal to accuracy and is given by:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \qquad (27)$$

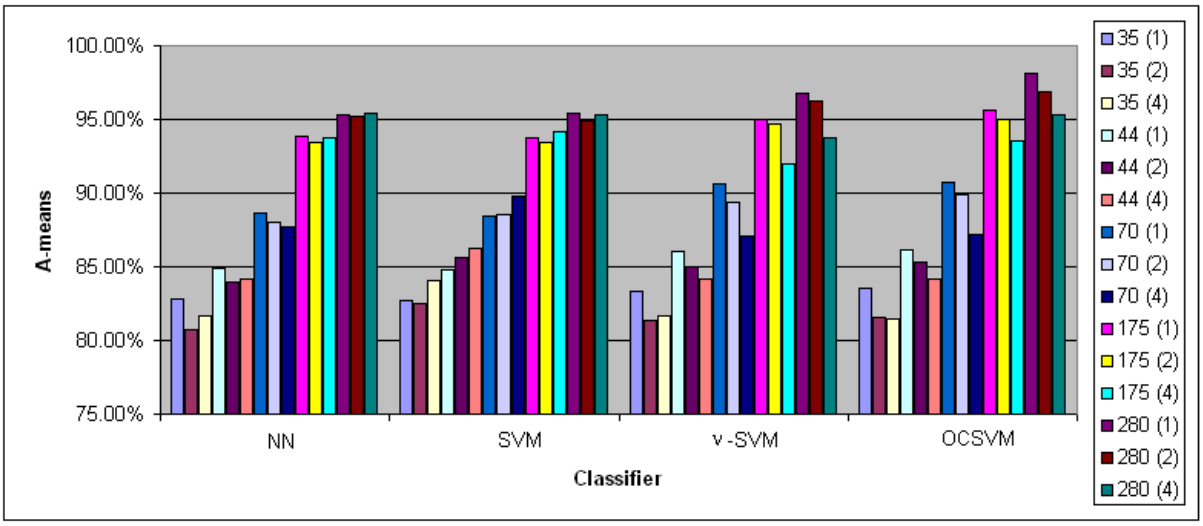


Fig. 8 Recognition performance obtained using standard LBP.

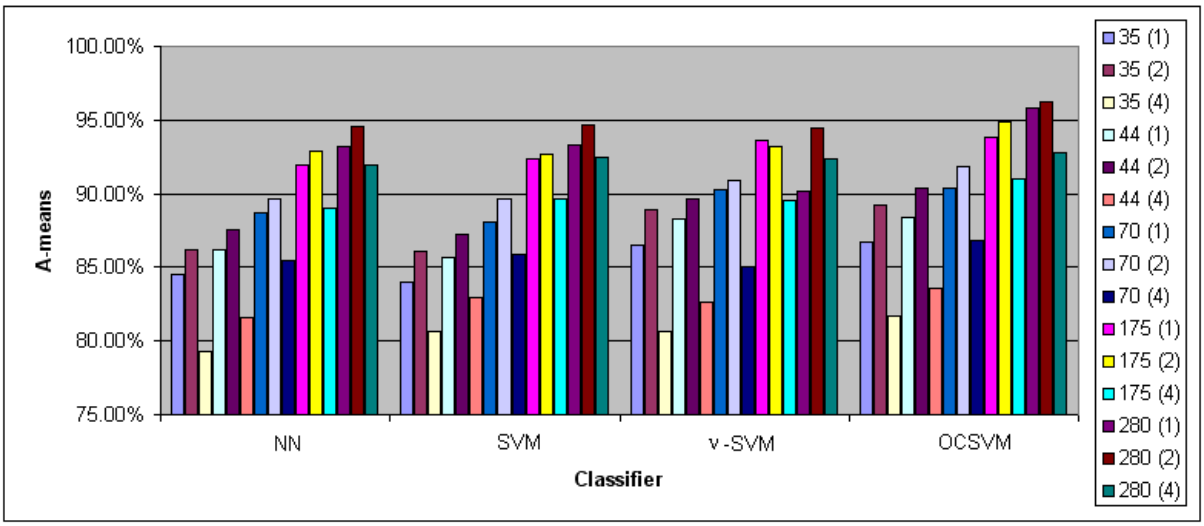


Fig. 9 Recognition performance obtained using Sobel-LBP.

The results obtained using the standard LBP are shown in Fig. 8 and Table 2 while the results obtained using the Sobel-LBP are shown in Fig. 9 and Table 3. The number inside a parenthesis and following the number of training data sets denotes the downsampling factor used. Fig. 8 and Fig. 9 show the recognition accuracy of the 4 classifiers, that is, nearest-neighbor (NN), SVM, v-SVM, and OCSVM. The number of bins used for the standard LBP is 256 while the Sobel-LBP consists of concatenated 256 bins for each horizontal and vertical filtered result, respectively. The LBP computation involves the use of overlapping blocks and only ROI pixels are considered as the central pixels in a block. The classifier parameters used for SVM are error weights C and radial basis function σ. For v-SVM and OCSVM, the parameters used are user-defined parameter v, RBF kernel width γ, and margin size ρ. These parameters were determined empirically. Overall speaking, the OCSVM classifier gives the best recognition performance according to the experimental results reported here. In general, many factors will affect the recognition accuracy, including feature selected, number of runs, classifier parameters, and the training dataset used.

For 10-runs of 5-fold cross validation using 80% of the training dataset, the combination of image downsampling by 1, LBP, and OCSVM classifier gives the following results:

a. Average of a-means: 98.10% (The maximum a-means is 100.00%, the minimum a-means is 93.28%),
b. Average of accuracy: 98.08%,
c. Average of TP rate: 98.16%,
d. Average of TN rate: 98.03%,
e. Average of precision [15]: 97.07%,
f. Average of Kappa value [16]: 0.9598 (almost perfect inter-rater agreement),

g. Average of g-means [17]: 98.07%,
h. Average of F-measure [15]: 97.55%,
i. Average number of support vectors is 256.86.

In some cases, the recognition performance of Sobel-LBP is better than standard LBP for smaller training datasets. Some downsampling factors used can cause a significant drop on recognition performance. When the standard LBP method is used, one reason for such a drop is the homogeneity of image such that the values of the pixels in the thresholded neighborhood, as performed in Fig. 5 (b), are all 1 for neighbor pixels because the pixel values in all neighbor pixels are equal to the central pixel, resulting in a single LBP value of 255. The corresponding histogram of standard LBP would be biased to a peak at the value of 255. This would in turn decrease the discrimination nature of the features derived from the histogram. When the Sobel-LBP method is used, the drop in recognition accuracy is significant for the case of downsampling factor by 4 because the image gradients obtained by using Sobel operators vary significantly in this case.

TABLE 2
RECOGNITION PERFORMANCE OBTAINED USING LBP

# of training images	DS	NN	SVM	v-SVM	OCSVM
35	1	82.84%	82.66%	83.33%	**83.59%**
	2	80.76%	**82.48%**	81.39%	81.53%
	4	81.65%	**84.06%**	81.71%	81.48%
44	1	84.89%	84.84%	86.08%	**86.18%**
	2	83.93%	**85.58%**	85.04%	85.30%
	4	84.19%	**86.25%**	84.19%	84.14%
70	1	88.66%	88.43%	90.63%	**90.71%**
	2	87.99%	88.57%	89.43%	**89.87%**
	4	87.69%	**89.76%**	87.11%	87.19%
175	1	93.81%	93.75%	94.96%	**95.67%**
	2	93.49%	93.42%	94.68%	**95.03%**
	4	93.75%	**94.13%**	92.00%	93.55%
280	1	95.27%	95.45%	96.73%	**98.10%**
	2	95.23%	94.90%	96.27%	**96.85%**
	4	**95.39%**	95.29%	93.80%	95.35%

TABLE 3
RECOGNITION PERFORMANCE OBTAINED USING SOBEL-LBP

# of training images	DS	NN	SVM	v-SVM	OCSVM
35	1	84.48%	83.96%	86.49%	**86.74%**
	2	86.19%	86.10%	88.92%	**89.22%**
	4	79.31%	80.69%	80.62%	**81.65%**
44	1	86.23%	85.62%	88.31%	**88.34%**
	2	87.51%	87.28%	89.60%	**90.36%**
	4	81.59%	82.93%	82.63%	**83.61%**
70	1	88.73%	88.11%	90.32%	**90.41%**
	2	89.67%	89.64%	90.89%	**91.89%**
	4	85.47%	85.87%	85.01%	**86.83%**
175	1	91.95%	92.39%	93.61%	**93.87%**
	2	92.92%	92.70%	93.19%	**94.87%**
	4	88.98%	89.61%	89.54%	**90.97%**
280	1	93.19%	93.26%	90.15%	**95.85%**
	2	94.51%	94.68%	94.42%	**96.19%**
	4	91.97%	92.47%	92.37%	**92.82%**

The computational time for testing a single image was evaluated by a personal computer with the following specification: Hardware: 3.00 GHz Intel (R) Pentium (R) 4 CPU and 1.00 GB RAM; Software: MATLAB 7.0 and Windows XP operating system.

The running time results for standard LBP and Sobel-LBP are shown in Table 4 and 5, respectively. In a real application, the capsule endoscope takes two images per second. Thus, a testing time lower than 0.5 seconds per image is desirable for real-time applications. The downsampling factors of 2 and 4 make the computation more efficient. On the average, the OCSVM classifier gives the lowest computational time. The results obtained for all cases are desirable for practical applications because the computational time is less than 0.5 seconds per image.

TABLE 4
AVERAGE COMPUTATIONAL TIME PER IMAGE IN SECONDS USING LBP

# of training images	DS	Time (NN)	Time (SVM)	Time (v-SVM)	Time (OCSVM)
35	1	0.20	0.25	0.11	0.11
	2	0.14	0.06	0.05	0.05
	4	0.13	0.03	0.05	0.03
44	1	0.19	0.11	0.11	0.11
	2	0.14	0.08	0.05	0.05
	4	0.11	0.06	0.03	0.05
70	1	0.28	0.23	0.22	0.22
	2	0.16	0.13	0.11	0.05
	4	0.13	0.11	0.09	0.03
175	1	0.19	0.11	0.11	0.14
	2	0.14	0.06	0.06	0.06
	4	0.13	0.05	0.05	0.03
280	1	0.20	0.11	0.11	0.13
	2	0.11	0.05	0.05	0.05
	4	0.11	0.03	0.06	0.03

TABLE 5
AVERAGE COMPUTATIONAL TIME PER IMAGE IN SECONDS USING SOBEL-LBP

# of training images	DS	Time (NN)	Time (SVM)	Time (v-SVM)	Time (OCSVM)
35	1	0.30	0.20	0.20	0.20
	2	0.19	0.08	0.06	0.06
	4	0.17	0.03	0.05	0.03
44	1	0.28	0.20	0.19	0.20
	2	0.19	0.09	0.09	0.06
	4	0.16	0.05	0.05	0.05
70	1	0.33	0.19	0.20	0.19
	2	0.17	0.06	0.08	0.06
	4	0.16	0.05	0.03	0.05
175	1	0.34	0.36	0.23	0.25
	2	0.19	0.08	0.08	0.08
	4	0.16	0.03	0.05	0.05
280	1	0.33	0.20	0.22	0.19
	2	0.17	0.08	0.06	0.08
	4	0.17	0.05	0.06	0.03

Since training sets and test sets vary from run to run, every run may have different misclassified images. Some images are misclassified due to poor lighting. Some normal images are misclassified because of some floating objects, such as food or liquid. Some abnormal images are misclassified because they contain very little

(or not enough) white spot lesions.

6. CONCLUSION

This paper proposes an abnormality detection system for capsule endoscope images. The system is intended to identify the images with certain symptoms of gastrointestinal tract automatically. Experimental results carried out on 10-runs of 5-fold cross validation using 80% of the training dataset showed that the combination of image downsampling by 1, standard LBP, and the OCSVM classifier outperforms other methods. The resulting accuracy is 98.10% and the computational time in the testing phase is below 0.5 seconds per image, which is applicable for practical applications.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Dr. Yi-Jen Fang, a senior physician of Show Chwan Memorial Hospital, Taiwan, for annotating the images and also to Ivanna Kristianti Timotius, a graduate of master's degree in Chung Yuan Christian University, for providing the contribution of the MATLAB program dealing with Support Vector Machines used in our experimental section.

REFERENCES

[1] D. D. Adler and C. J. Gostout, "Wireless capsule endoscopy," *Hospital Physician*, pp. 14-22, May 2003.

[2] M. T. Coimbra and J. P. S. Cunha, "MPEG-7 visual descriptor – contribution s for automated feature extractor in capsule endoscopy," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 16, no. 5, pp. 628-637, May 2006.

[3] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proc. of Int. Workshop on Analysis and Modeling of Faces and Gestures*, pp. 168-182, 2007.

[4] M. A. Savelonas, D. K. Iakovidis, and D. E. Maroulis, "An LBP-based active contour algorithm for unsupervised texture segmentation," in *Proc. of the 18th Int. Conf. on Pattern Recognition*, vol. 2, pp.279-282, 2006.

[5] V. Vapnik, *Statistical Learning Theory*, Springer, Berlin, Hiedelberg, New York, 1998.

[6] I. K. Timotius, S. G. Miaou, and Y. H. Liu, "Abnormality detection for capsule endoscope images based on color histogram and support vector machines," in *Proc. of the 21st Conf. on Computer Vision, Graphics, and Image Processing*, Yilan, Taiwan, 2008.

[7] P. H. Chen, C. J. Lin, and B. Schölkopf, "A tutorial on ν-support vector machines," *Applied Stochastic Models in Business and Industry*, vol. 21, pp. 111-136, 2005.

[8] A. Takeda and M. Sugiyama, "ν-support vector machine as conditional value-at-risk minimization," in *Proc. of the 25th Int. Conf. on Machine Learning*, Helsinki, Finland, 2008.

[9] L. Zhuang and H. Dai, "Parameter optimization of kernel-based one-class classifier on imbalance learning," *J. of Computers*, vol. 1, no. 7, pp. 32-40, Oct./Nov. 2006.

[10] L. M. Menevitz and M. Yousef, "One-class SVMs for document classification," *J. of Machine Learning Research*, vol. 2, pp. 139-154, 2001.

[11] S. Haykin, *Neural Network: A Comprehensive Foundation*, Prentice-Hall, New Jersey, 1999.

[12] S. Zhao, Y. Gao, and B. Zhang, "Sobel-LBP," in *Proc. of IEEE Int. Conf. on Image Processing*, pp. 2144-2147, 2008.

[13] C. C. Chang and C. J. Lin, *LIBSVM: A Library for Support Vector Machines*, 2001.

[14] C. H. Lin, J. C. Liu, and C. H. Ho, "Anomaly detection using LibSVM training tools," in *Proc. of Int. Conf. on Information Security and Assurance*, pp. 166-171, 2008.

[15] T. Z. Tan, G. S. Ng, and C. Quek, "Complementary learning fuzzy neural network: an approach to imbalanced dataset," in *Proc. of Int. Joint Conf. on Neural Networks*, pp. 2306-2311, Aug. 2007.

[16] A. Ben-David, "What's wrong with the hit ratio?," *IEEE Intelligent Systems*, vol. 21, issue 6, pp. 68-70, Nov.-Dec. 2006.

[17] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training datasets: one-sided selection," in *Proc. of the 14th Int. Conf. on Machine Learning*, pp. 179-186, 1997.

[18] I. K. Timotius, *Abnormality detection for capsule endoscope images based on support vector machines*, Master Thesis, Chung Yuan Christian Univ., Chungli, Taiwan, Jan. 2009.

[19] X. Song, G. Cherian, and G. Fan, "A *v*-insensitive approach for compliance monitoring of the conservation reserve program," *IEEE Trans. on Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 99-103, April 2005.

[20] K. A. Heller, K. M. Svore, A. D. Keromytis, S. J. Stolfo, "One class support vector machines for detecting anomalous windows registry accesses," *in Proc. of the ICDM Workshop on Data Mining for Computer Security (DMSEC)*, Melbourne, Florida, Nov 2003.