Infected Cell Identification in Thin Blood Images Based on Color Pixel Classification: Comparison and Analysis

Gloria Díaz, Fabio Gonzalez, and Eduardo Romero

Bioingenium Research Group, National University of Colombia, Bogotá, Colombia
{gmdiazc,fagonzalezo,edromero}@unal.edu.co
http://www.bioingenium.unal.edu.co

Abstract. Malaria is an infectious disease which is mainly diagnosed by visual microscopical evaluation of Giemsa-stained thin blood films using a differential analysis of color features. This paper presents the evaluation of a color segmentation technique, based on standard supervised classification algorithms. The whole approach uses a general purpose classifier, which is parameterized and adapted to the problem of separating image pixels into three different classes: parasite, blood red cells and background. Assessment included not only four different supervised classification techniques - KNN, Naive Bayes, SVM and MLP - but different color spaces -RGB, normalized RGB, HSV and YCbCr-. Results show better performance for the KNN classifiers along with an improving feature characterization in the normalized RGB color space.

Keywords: Cell detection, Supervised classification, Color spaces, Performance comparison.

1 Introduction

Malaria is a leading cause of morbidity and mortality in tropical and sub-tropical countries, with an estimated of 300 to 500 worldwide million infections per year and 1 to 2 million deaths [1]. Plasmodium falciparum is the most mortal of the four species. In recent years, many research works have been addressed to development of new therapeutic alternatives for control of this disease [2], which involves in vitro drug susceptibility analysis by parasitism level quantification. Although different approaches have been developed for determining the level of infected erythrocytes with Plasmodium falciparum, visual microscopical evaluation of Giemsa-stained thin blood films is so far the most widely used in development countries. Its main drawback is that it is a subjective and time consuming method which demands trained technical personnel. In this context, development of mechanisms that automate the process of evaluation and quantification in thin blood films, becomes a high priority.

Several digital image processing techniques have been previously used for detecting malaria parasites on Giemsa stained slides [3], [4], [5], [6], [7].Plasmodium

L. Rueda, D. Mery, and J. Kittler (Eds.): CIARP 2007, LNCS 4756, pp. 812–821, 2007.

falciparum parasites are highlighted in a dark purple colour, while erythrocytes are colored in slight pink colors. Object detection has been performed using a threshold on single components of the RGB and HSV histograms [3], [6]. Likewise, parasite detection has been achieved in two consecutive steps: a former stained/non-stained pixel classification - based on the RGB values - is followed by setting the pixel to any of the parasite/non-parasite categories - based on other features such as shape, color and Hu moments [4]. Finally, the color co-occurrence matrix has been calculated for pixel classification in cells previously detected by a template matching strategy [5].

In this paper, we present a very simple approach for automatic identification of infected and no infected erythrocytes in thin blood images by means of a supervised pixel classification method. Herein, an exhaustive study of the effect of selecting both a color space representation and a particular classifier on actual Plasmodium falciparum slides is presented. We investigated four color spaces (RGB, normalized RGB, HSV and YCbCr) and four supervised classification algorithms (Naive Bayes, SVM, KNN and Neural network). A separate analysis was performed only on the chrominance component of each color space. This paper is organized as follows: color representations and color pixel classification algorithms are described in Section 2, comparison results are presented in section 3 and discussion and conclusions are given in Section 4.

2 Cell Identification Based on Pixel Classification

The overall approach for identification of cells proposed in this paper is illustrated in figure 1. First, a set of training samples was manually extracted by an expert. Each training sample corresponded to a pixel labeled as erythrocyte, parasite or background. Then, a classification model was trained using these sample pixels, which was used for classifying the whole color space (RGB, normalized RGB, HSV or YCbCr). The classified color space was so used as a look-up table (LUT) for classifying pixels. Finally, the image was re-colored in three gray level values (background, erythrocyte and parasite) and a a two-scan connected component labeling algorithm [8] was applied for identifying and counting the objects present in the image.

2.1 Color Representation

As mentioned before, different color spaces were used to building the pixel classification model since features are differently represented in each. For instance, Di Ruberto found found that it is easier to identify parasites in the S component of the HSV color space [3]. The different color spaces assessed in this work are described in the following subsections.

RGB. This color space is used for acquiring and displaying color digital images. Each color pixel is represented by its three components: R(Red), G(Green) and B(Blue).

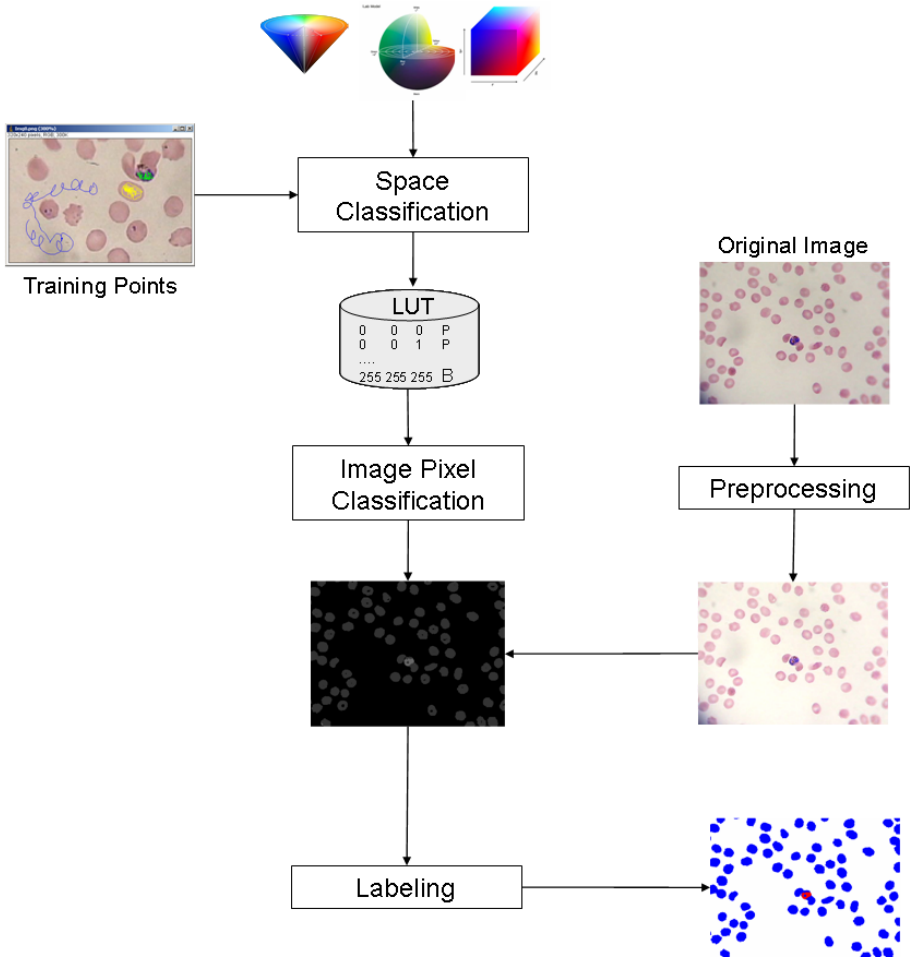

Fig. 1. The main steps of the whole erythrocytes and parasites detection

Normalized RGB (RGBn). This transformation is obtained by a simple procedure of RGB normalization:

$$\begin{aligned} r_N &= \frac{R}{R+G+B} \\ g_N &= \frac{G}{R+G+B} \\ b_N &= \frac{B}{R+G+B} \end{aligned} \tag{1}$$

This nonlinear transformation reduces the sensitivity of the distribution to color variability, making it more robust to illumination changes than RGB. Since $r + g + b = 1$, when two components are given, the third component can be determined. Thus, only two of these three were used.

HSV. This transformation decorraleted color and intensity information in the image. Color information is represented by hue and saturation components, while intensity is determined by the value component. Hue defines the basic color in the pixel, saturation measures its colorfulness related to its brightness and value corresponds to the luminance color.

YCbCr. Is a family of color spaces commonly used to represent digital video. Luminance information is stored as a single component (Y), and chrominance correspond to the two color-difference components (Cb and Cr). We have used the YCbCr transformation specified in the ITU-R BT.601 standard for computer-display oriented applications.

2.2 Classification Models

Supervised learning is the area of machine learning or pattern recognition, that addresses the problem of building models for performing classification or regression tasks. This is one of the areas more deeply and extensively studied in machine learning. Tens of algorithms have been proposed, ranging from biologically inspired to pure statistical techniques. Each has its own weaknesses and strengths and, according to the No-Free Lunch Theorem [9], there is not one that could be deemed as superior to the rest for any classification task. In general, one algorithm may outperform another algorithm in a particular task, but may under perform in other task. According to the previous discussion different algorithms were tried. The chosen algorithms are representative of the state of art and of different approaches to supervised learning.

The Naive Bayes Approach. Likely, this is the simpler classifier and is based on the hypothesis that features are conditionally independent, which in terms of the Bayes theorem amounts to

$$P(C|x_1, x_2, \ldots, x_n) = \frac{1}{K} P(C) \prod_{i=1}^{n} P(x_i|C) \tag{2}$$

where K is a constant dependent only on x_i and $P(c)$ is a prior probability of the class C, which is herein calculated during the training phase by merely counting the number of occurrences in the training data set.

The *k*-NN decision rule. The k-nearest neighbors method is well known used in the pixel classification problems [10][11]. It is an intuitive method that classifies unlabeled samples based on their similarity with samples in the training set. Given the knowledge of N prototype features (vectors of dimension Σ) and their correct classification into M classes, the k-NN rule assigns an unclassified pattern to the class that is most heavily represented among its k neighbors in the pattern space (under some appropriate metric).

A Neural Network strategy (MLP). Networks with organizations that emulate nervous system connections have been used in a large variety of image

segmentation problems. Herein, a Multi Layer Perceptron (MLP) trained using back-propagation was used [12]. The idea of this method is to connect layers of "neurons", while the particular neuron response is modeled with a continuous sigmoid approximation .

The *SVM* algorithm. A support vector machine (SVM) is a classification model that finds an optimal separating hyperplane that discriminates two classes. In principle, a SVM is a linear discriminator, however it can perform non-linear discrimination thanks to the fact that it is a kernel method. In this work, a version of SVM that uses sequential minimal optimization algorithm is used [13]. The multi-class classification problem is solved creating one classifier for each pair of the target classes, ignoring instances that belong to other classes and estimating a probability for each target class. Absolute probability estimate for each class is computed combining the probability estimate from all pairwise classifiers.

3 Experimentation

3.1 Data Set

A total of 25 microscopical fields from three different thin blood films were digitized using a Sony high resolution digital video camera Handycam DCR-HC85 (640 × 480 pixels to 1200 × 16000), coupled to a Carl Zeiss Axiostar Plus microscope, provided with Carl Zeiss 426126 and 456006 adapters (Carl Zeiss, Light Microscopy, Gottingen, Germany). Use of intermediate lens and a ×100 power objective yielded a total of ×1006 magnification. Optical image was a $102 \times 76 \mu m^2$ for a 640 × 480 image size, resulting in a total resolution of $0.0252\, \mu m^2/pixel$. A total of 1226 erythrocytes and 60 parasites were found in these images, indicating that the most relevant class (parasite) was barely represented by a 5 %.

Before applying the classification process, a correction of the luminance differences in the original image is performed through a local low pass filter. This filter is essentially a local adaptive filter, defined for a window size of the larger image feature, i.e. a typical erythrocyte size. Firstly, the $m \times n$ RGB luminance and chrominance image components are decorrelated through a YCbCr transformation. Luminance channel is split into disjoint regions of approximately the larger feature in the image and a mean pixel value is calculated from each. These mean values make up a matrix which is smoothed out using a moving smaller window, whose size is adjusted in order to eliminate the tiling effect of the filter. Afterward, luminance is corrected by ruling out the lower frequencies found before. Finally, the color image is re-constructed using this luminance correction and the original chrominance information. Figure 2 displays two microscopical images obtained from thin blood smears. Upper row displays the original digital images, while bottom row shows the obtained images using the proposed filter.

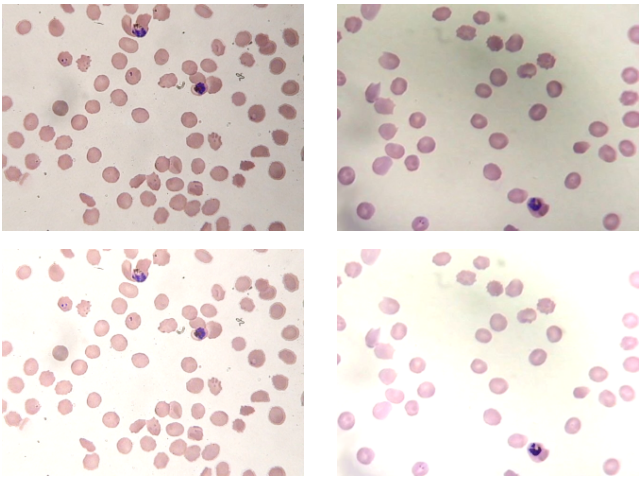

Fig. 2. Pre-processed image results. Fist row corresponds to the original microscopical images, bottom row displays filtered images obtained from the original ones.

3.2 Experimental Setup

Four classification algorithms were assessed (Naive Bayes, KNN, SVM and MLP). Each classification model was tuned independently for its own particular set of parameters as follows: k-NN was assessed by varying the k odd nearest neighbors between 1 and 15. *SVM* was evaluated with different kernels i.e. radial or polynomial [13]. In both cases C gap was set to 1. Additionally, in the former case the γ parameter was varied between 100 and 1000 with increment steps of 100, while an optimal polynomial degree was determined for the later case $(1, 2, 3)$. The Bayes algorithm was trained with a normal distribution and a 95 % confidence interval. The neural network was provided with a hidden layer on which the number of neurons was varied between 3 and 9, with increment steps of 3 and error rates among $(0.005, 0.1, 0.2, 0.5, 0.9)$. All classifiers were trained using sets of 500, 1000 and 2000 pixel samples, classified manually by an expert in two representative images.

The different color spaces stand for the characteristic spaces so that six different feature vectors were analyzed : four complete (RGB, HSV, YCbCr, normalized RGB) and two incomplete color spaces (HS, CbCr).

3.3 Evaluation

Classification performance was assessed based on a reference-manual segmentation, using two strategies: pixel classification and interest-object detection (erythrocytes and parasites) rates. In this analysis, performance estimation through a conventional accuracy comparison results inappropriate because of the high imbalanced class distribution of parasites related to erythrocytes and background. That is to say, the assessment may report a high accuracy even if parasites are

not identified. As an alternative to accuracy, we used the F-measure, or effectiveness measure, [14] computed as $F_\beta = \frac{(1+\beta)*RC*PR}{\beta*PR+RC}$, with RC (recall) defined as $\frac{TP}{TP+FN}$ and PR (precision rate) computed as $\frac{TP}{TP+FP}$, where TP stands for the true positives, FN for the false negatives and FP for the false positives. The $\beta > 0$ coefficient controls the relative importance of recall and precision rates; $\beta = 1$ gives the same importance to both measures, whilst precision rate is more important with a higher value of β. Herein, β was set to $2/3$, since we attempt to detect as many objects as possible even at expense of lower precision. $F_\beta = 1$ means a perfect score, i.e. $PR = RC = 1$.

Pixel wise evaluation is performed by comparing re-colored images, generated by each set of training points and classifier parameters, with a manual segmentation. In this case, the test set is composed of labeled pixels. Precision and recall are calculated based on the number of rightly/wrongly classified pixels.

Interest-object wise evaluation is based on the objects identified by a method that includes the classification process as the first step. After the application of the classification method, objects are identified as follows: a basic filtering process is performed for keeping only relevant objects in the image; thus, small or large regions identified as flaws (particulate matter from the stain or from fragments of released hemazoin, acquisition artifacts) are removed; likewise, near unconnected segments are evaluated for establishing their relevance to a given parasite, if they are relevant they are considered as a unique object.

For interest-object wise evaluation, the test set is composed of images where interest-objects (erythrocytes and parasites) are labeled. Precision and recall are calculated based on the rightly/wrongly identified objects.

3.4 Results and Discussion

As it was mentioned before, training sets with different sizes were used. It was noticed that, in all the cases, increasing the size of the training set from 500 to 2000 did not improve significantly the performance of the different classifiers. Therefore, all subsequent experiments were performed using a 500 elements training data set.

The different classification algorithms were trained and evaluated with different parameter values, as mentioned in the experimental setup (for the sake of brevity, these intermediate results are not shown). The best parameter values for each algorithm were identified: K = 15 for KNN, error rate = 0.1 and 6 neurons in hidden layer for MLP, polynomial degree = 1 for SVM with polynomial kernel (SVM_P) and $\gamma = 100$ for SVM with radial base kernel.

Effectiveness measure results for pixel wise and interest-object wise evaluations are shown in Tables 1 and 2. The F-measure is reported independently for the erythrocyte (F_β^{Er}) and parasite (F_β^P) classes, respectively.

Pixel Wise Evaluation. Pixel wise evaluation results are shown in Table 1. F_β values suggest that performance is good for the erythrocyte class, while it is not as good for the parasite class. The pattern is the same for all classifier algorithms and color spaces. Our hypothesis is that the complex mix of colors,

present in the parasites, makes it difficult to discriminate individual pixels using only color information.

The best overall performance is accomplished by the combination of a KNN classifier and YCbCr color space, with $F_\beta^{Er} = 0.95$ and $F_\beta^P = 0.72$. However, there are other combination that produce similar results such as KNN classifier and normalized RGB color space ($F_\beta^{Er} = 0.95$ and $F_\beta^P = 0.71$), MLP-classifier and normalized RGB color space ($F_\beta^{Er} = 0.94$ and $F_\beta^P = 0.71$) and SVM_P-classifier and YCbCr color space ($F_\beta^{Er} = 0.94$ and $F_\beta^P = 0.72$).

From the point of view of color space, normalized RGB and YCbCr have better performance. This indicates that these color spaces emphasize the differences between classes. From a classifier stantpoint KNN, MLP and SVM_P clearly outperform SVM_{RBF} and Naive Bayes.

Table 1. F_β measure results for pixel wise evaluation for different classification algorithms and color spaces. F-measure is reported independently for the erythrocyte (F_β^{Er}) and parasite (F_β^P) classes.

	Naive Bayes		KNN $k = 15$		MLP $er = 0.1, n = 6$		SVM_P $degree = 1$		SVM_{RBF} $\gamma = 100$	
ColorSpace	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P
RGB	0.88	0.13	0.95	0.69	0.95	0.68	0.94	0.68	0.92	0.17
HSV	0.90	0.35	0.89	0.70	0.90	0.68	0.89	0.65	0.89	0.05
HS	0.90	0.37	0.91	0.72	0.90	0.62	0.89	0.67	0.94	0.62
RGBn	0.95	0.52	**0.95**	**0.71**	**0.94**	**0.71**	0.94	0.70	0.93	0.45
YCbCr	0.89	0.52	**0.95**	**0.72**	0.94	0.68	**0.94**	**0.72**	0.86	0.07
CrCb	0.84	0.68	0.85	0.71	0.86	0.63	0.86	0.70	0.86	0.50

Interest-Object Wise Evaluation. Table 2 shows the results of the interest object wise evaluation. The best overall performance is clearly accomplished by the combination of a KNN-classifier and normalized RGB color space with $F_\beta^{Er} = 0.99$ and $F_\beta^P = 0.83$, followed by SVM_P-classifier and YCrCb color spaces ($F_\beta^{Er} = 0.97$ and $F_\beta^P = 0.81$). This means that normalized RGB and YCrCb color spaces produced again the best results. This is no really surprising as object identification is based on pixel classification.

With regard to the classifier algorithm, SVM_{RBF} and Naive Bayes produced the best results for erythrocyte detection, however their performance on parasite detection was really poor. Both KNN and SVM_P produced a good balance of parasite and erythrocyte detection.

An interesting finding in these results, is the fact that the performance is much better at the level of object identification than at the level of pixel classification. The main reasons is that pixel classification is more sensitive to noise, while the

Table 2. F_β measure results for erythrocytes and parasites detection. for different classification algorithms and color spaces. F-measure is reported independently for the erythrocyte (F_β^{Er}) and parasite (F_β^P) classes.

	Naive Bayes		*KNN* $k=15$		*MLP* $er=0.1, n=6$		SVM_P $degree=1$		*SVM_RBF* $\gamma=100$	
ColorSpace	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P	F_β^{Er}	F_β^P
RGB	0,98	0,19	0,99	0,76	0,99	0,78	0,93	0,73	0,99	0,18
HSV	0,97	0,43	0,96	0,75	0,97	0,72	0,96	0,71	0,98	0,11
HS	0,97	0,44	0,97	0,75	0,97	0,65	0,95	0,74	0,99	0,67
RGBn	0,99	0,57	**0,99**	**0,83**	0,98	0,78	0,96	0,74	0,98	0,40
YCrCb	0,97	0,47	0,97	0,77	0,98	0,77	**0,97**	**0,81**	0,98	0,10
CrCb	0,95	0,80	0,92	0,81	0,94	0,72	0,90	0,76	0,95	0,51

object identification process is able to eliminate this noise thanks to the filtering process that improves the results of the pixel classification step.

4 Conclusions

A simple an efficient method for parasite and erythrocyte detection in thin blood images was proposed. The approach is based on a classification process that finds boundaries that optimally separate a given color space in three classes, namely, background, erythrocyte and parasite. The classified color space is stored and used as a look-up table for classifying pixels from new images.

The method was evaluated at two levels: pixel classification and object detection. Different classification algorithms and color spaces were evaluated. KNN algorithm with normalized RGB color space was found to have higher detection performance compared to other tested classifiers. Furthermore, this color space requires less computational resources as only two components are needed to fully determine a point in this space (the third one is calculated from the first two). Color spaces traditionally used as RGB or HSV produced poorer results. The performance result at the object-identification level was superior to the performance at the pixel-classification level. This shows that the filtering step used by the object-identification process is able to reduce noise, making the overall process robuster.

Future work is focused on evaluating the feasibility of using combined color spaces and more specialized classification algorithms.

Acknowledgments

This work was partially supported by a grant from the Colombian Institute for the Advancement of Science and Technology (COLCIENCIAS), Grant no.

109-2005. Smears used in this study were supplied by the Research Group in Bioactive Principles in Medicinal Plants from National University of Colombia.

References

1. WMR, UNICEF: World malaria report. Technical report, WMR and UNICEF (2005)
2. OPS: The health in the americas. Technical Report 1, Pan-american organization of the Health (1998)
3. di Ruberto, C., Dempster, A., Khan, S., Jarra, B.: Analysis of infected blood cell images using morphological operators. Image and Vision Computing 20(2), 133–146 (2002)
4. Tek, F., Dempster, A., Kale, I.: Malaria parasite detection in peripheral blood images. In: Proceeding of British Machine Vision Conference (2006)
5. Halim, S., Bretschneider, T.R., Li, Y., Preiser, P.R., Kuss, C.: Estimating malaria parasitaemia from blood smear images. In: Proceedings of the IEEE International Conference on Control, Automation, Robotics and Vision (2006)
6. Ross, N.E., Pritchard, C.J., Rubin, D.M., Dus, A.G.: Automated image processing method for the diagnosis and classification of malaria on thin blood smears. Medical and Biological Engineering and Computing 44, 427–436 (2006)
7. Sio, S.W.S., Sun, W., Kumar, S., Bin, W.Z., Tan, S.S., Ong, S.H., Kikuchi, H., Oshima, Y., Tan, K.S.W.: Malariacount: an image analysis-based program for the accurate determination of parasitemia. Microbiological Methods 68 (2007)
8. di Stefano, L., Bulgarelli, A.: A simple and efficient connected components labeling algorithm. In: Proceedings of the 10th International Conference on Image Analysis and Processing (1999)
9. Wolpert, D., Macready, W.: No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation 1, 67–82 (1997)
10. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Transactions on Information Theory IT-13, 21–27 (1967)
11. Fix, E., Hodges, J.: Discriminatory analysis, non-parametric discrimination. Technical Report Project 21-49-004, Rept. 4, Contract AF41(128)-131, USAF School of Aviation Medicine, Randolf Field, Texas (February 1951)
12. Rumelhart, D., Hinton, G., Williams, R.: Parallel Distributed Processing: Explorations in Macrostructure of cognition, vol. I. Badford Books, Cambridge. MA (1986)
13. Platt, J.: Machines using sequential minimal optimization. In: Schoelkopf, B., Burges, C., Smola, A. (eds.) Advances in Kernel Methods - Support Vector Learning, MIT Press, Cambridge (1998)
14. Daskalaki, S., Kopanas, I., Avouris, N.: Evaluation of classifiers for an uneven class distribution problem. Applied Artificial Intelligence 20, 381–417 (2006)