Machine Vision and Applications
DOI 10.1007/s00138-010-0275-y

ORIGINAL PAPER

The application of support vector machine classification to detect cell nuclei for automated microscopy

Ji Wan Han · Toby P. Breckon · David A. Randell · Gabriel Landini

Received: 8 September 2009 / Revised: 11 May 2010 / Accepted: 12 May 2010

Abstract The detection of cell nuclei for diagnostic purposes is an important aspect of many medical laboratory examinations. Precise location of cell nuclei can aid in correct diagnosis and aid in automated microscopy applications, such as cell counting and tissue architecture analysis. In this paper, we investigate the use of support vector machine classification based on Laplace edge features for this task. Compared with existing applications, we used only one type of cell nucleus images to train the classifier but this classifier can locate other two types of cell nuclei with different stains and scales successfully. The results illustrate that such a data driven approach has remarkable detection and generalization performance.

Keywords Cell nuclei detection · Automated microscopy · Support vector machines

J. W. Han (✉) · T. P. Breckon
School of Engineering, Cranfield University,
Cranfield MK43 0AL, UK
e-mail: j.w.han@cranfield.ac.uk

T. P. Breckon
e-mail: toby.breckon@cranfield.ac.uk

D. A. Randell · G. Landini
Oral Pathology Unit, School of Dentistry,
University of Birmingham, St. Chad's Queensway,
Birmingham B4 6NN, UK
e-mail: d.a.randell@bham.ac.uk

G. Landini
e-mail: g.landini@bham.ac.uk

1 Introduction

The cell is the smallest unit of living organisms, and is often referred to as a building brick of life [1]. The analyses of individual cells and cells in groups (tissues) often indicate useful information about living organisms, physiological and diseased states and the response to therapies. The first step of automating such diagnosis and analysis is the detection of individual cells. The cell nucleus is the most conspicuous organelle found in a eukaryotic cell [1] and thus their detection facilitates the task of locating cells in a sample.

For lesions that cannot be identified clinically, small tissue samples (biopsies) are taken and submitted to histopathological (microscopic) examination. This consists of visual analysis performed by expert observers (histopathologists) who interpret morphological changes in those cells and tissues. However, such an approach introduces an unavoidable subjective element to histopathological diagnosis because simple visual observation is unable to quantify the extent of those morphological features. Consequently, the main aspect of histopathological diagnosis remains prone to individual differences in the perceptual abilities of the observers. This, in turn, introduces uncertainties in reproducibility (both from the point of view of intra- and inter- observer variations) and precludes the use of statistical (and consequently) evidence-based knowledge to inform disease progression and treatment evaluation.

Numerous approaches have been investigated to achieve stable, objective and generalized detection of cell nuclei. Traditional image processing methods have had an extensive application in this domain. Landini [14] analysed epithelial lining architecture in radicular cysts and odontogenic keratocysts applying image processing algorithms to follow a traditional cell isolation based approach. In this application, the watershed method is applied with the help of morphological

Published online: 05 June 2010

methods to achieve the segmentation of individual cell nuclei. This formed the basis for later estimation of tissue layer level and architectural analysis of oral epithelia [15]. Kruk et al. [13] applied morphological operations and the watershed algorithm to extract individual cells from colon tissue. These cells were then fed into neural network classifiers for subsequent recognition of colon cell types achieving a mean discrepancy rate of 6%. Schnorrenberg et al. [25] developed a computer-aided detection system for tissue cell nuclei in histological sections using adaptive thresholding to locate cells. Detection and classification of individual nuclei as well as biopsy grading performance was shown to be promising as compared to that of two experts. Sensitivity and positive predictive value were measured to be 83 and 67.4%, respectively.

Recently, with the rapid development of machine learning algorithms, some medical image processing applications have adopted further techniques in order to improve detection and classification performance. Among these algorithms, Support Vector Machines (SVMs) are broadly known and have shown a high classification performance on many applications, including cell yeast cells on suspension in bioreactors [31], blood cell sorting and tissue cells [28], cells in culture using fluorescent microscopy [29] and on sections of brain tumors [6]. Wei et al. [31] applied two support vector machine based classifiers to separate cells from background, and to distinguish live from dead cells afterwards. These classifiers displayed high accuracy and stability. They [30] also developed a machine vision system based on a supervised learning technique that is able to learn from images of cell populations and trains a number of classifiers. They subsequently employed a SVM classifier to determine the viability of each tested cell. Rahman et al. [20] applied SVMs to medical image annotation and retrieval achieving an automatic image annotation accuracy of 56.7% and a retrieval accuracy of 72.96%. El-Naqa et al. [5] applied SVMs to detect microcalcifications in mammogram images, which outperformed all the other methods considered within the study. A sensitivity as high as 94% was achieved by the SVM method at an error rate of one false-positive cluster per image. Other algorithms also displayed remarkable performance. Nattkemper et al. [18] designed a neural cell detection system (NCDS) for the automatic quantization of fluorescent lymphocytes in tissue sections. This system acquired visual knowledge from a set of training cell-image patches selected by a user. The trained system evaluated an image calculating the number, the positions and the phenotypes of the fluorescent cells. The NCDS detected a minimum of 95% of the cells. In our prior cell nuclei detection work [7], a machine learning based approach (cascaded Haar classification [27]) was used to explore the feasibility of using such algorithms to classify three types of cystic lesions of the jaws: solitary odontogenic keratocysts, basal cell naevus syndrome associated odontogenic keratocysts, and radicular cysts [14,15,26]. In those experiments a total correct classification rate of 86% was achieved. That method showed successful detection of individual cell nuclei within the pathological slides in addition to promising classification rates on the cyst subtypes [7].

Despite the growing number of papers on cytology and image analysis, it is necessary to develop methods of detection and classification that can be applied to standardised sample preparation modalities for the ultimate goal of high throughput sample screening. In this paper, we investigate the possibility of SVMs detection of nuclei on stained monolayer cell cultures.

2 Using support vector machine classifiers

2.1 Support vector machines

SVMs, based on the principle of structural risk minimization now form a well established approach in the application of machine learning algorithms and are proving to be particularly promising when used to construct accurate models based on large feature spaces [4,9,17]. Particularly, SVMs deliver state-of-the-art performance in real-world applications [4,5,8–10,19,31]. They have some superiorities over other approaches, especially: (a) global minimum solution, and (b) learning and generalization in huge dimensional input spaces [4,9]. Essentially, they use a hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. The aim is to find a hyperplane which can classify two classes of data correctly, by maximizing the distance between the two classes of data and the hyperplane, in a space of higher dimension. From Fig. 1, we can see class 1 and class 2 can be separated by many hyperplanes but only the optimal hyperplane separates two classes with maximum margin. Margin b and c are shorter than margin a. Those points lying on the margin generated by the optimal hyperplane are support vectors.

SVMs [2] perform pattern classification by determining the separating hyperplane at a maximum distance to the closest points in the training set. These points are called support vectors. The decision function of the SVM has the form:

$$f(x) = \text{sign}\left[\sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b\right], \tag{1}$$

Where x is the data point to be classified, x_i are support vectors, N is the number of support vectors, b is a constant decided from training and $y_i \in \{-1, 1\}$ is the class label of the support vector x_i. The coefficients α_i are the solution of a quadratic programming problem. The margin, which is

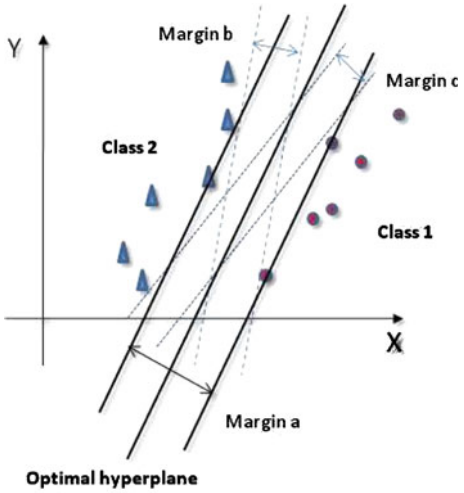

Fig. 1 SVM finds the hyperplane separating two classes with maximum distance

Table 1 The most popular kernel types

Kernel type	Kernel
Linear	$x_i^T x$
Polynomial	$(x_i^T x + \tau)^d$
Radial basis function	$exp(-\frac{\|\|x-x_i\|\|^2}{2\sigma^2})$
Sigmoid	$tanh(\kappa x_i^T x + \theta)$

the distance of the support vectors to the hyperplane, is thus given by:

$$M = \frac{1}{\sqrt{\sum_i^N \alpha_i}} \tag{2}$$

The margin is an indicator of the separability of the data within the dimensionality of the hyperplane (Fig. 1).

The kernel function plays a key role for SVMs in solving real-world problems because many such applications are not linearly separable in their original dimensional space (i.e. that of the input). By applying a kernel transform K, the input data vectors are mapped into a higher-dimensional space. In this space, the mapped data vectors could be linearly separable or have improved separability [4]. There are several popular kernel transforms shown in Table 1. The Radial Basis Function (RBF) kernel is commonly considered as the most powerful but linear kernels are best understood and are the simplest to apply [4]. By doing some parameter configuration, a RBF kernel can be converted to a linear one [11]. The linear kernel is suitable for solving linear separable problems. If target classes are not linear separable, they have to be projected to a higher dimension space where they are linear separable or easier to be separated. This process is done using non-linear kernels, such as the RBF kernel. Figure 2 depicts an exemplar classification problem in 2D together with the application of

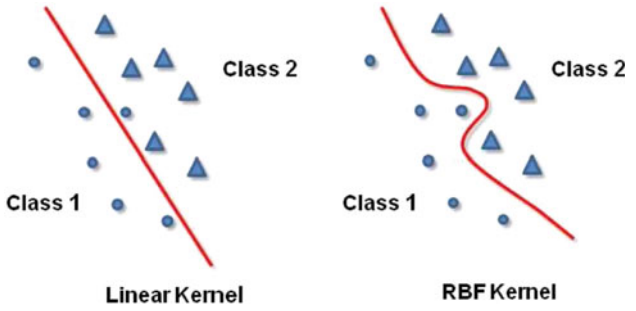

Fig. 2 Classification using linear kernel and RBF (radial basis function) kernel. Linear kernel and RBF kernel have different classification performance

linear and RBF kernels. In this example, a linear kernel cannot separate two classes without misclassification; a class 1 point is classified as class 2. In contrast, a RBF kernel separates them correctly. The selection of a suitable kernel type for a SVM classifier applied to a given problem will be discussed in Sect. 2.2.

2.2 Training support vector machines

A SVM classifier has to be initially trained before use for classification. During training, samples of two classes are presented using a training procedure. After analysing these training data, the classifier finds the hyperplane which separates two classes with maximum margin in a given dimensional space using a given kernel function, K. The aim of training the SVM classifier is to find a suitable kernel, K and its associated parameters. A commonly used parameter evaluation method is cross-validation [12]. This is a statistical method of testing hypotheses that keeps a subset of data as a test set and the remaining data as a training set. There are several commonplace cross-validation methods with *k-fold* cross-validation being the most popular [16]. This procedure divides original data into k subsets and keeps one of them as a test set while the remaining $k-1$ subsets are used to train the classifier. This process is repeated until every subset has been used as test set and the mean correct rate is taken as the single estimation of the classifier. Cross-validation helps to eliminate the unilateral testing hypotheses suggested by the data [16] (i.e. over-fitting), so that it can give a comprehensive estimation of parameter configuration.

In order to utilize the cross-validation method to find the best kernel, K and its parameters, a grid search method is introduced [3,11]. The grid search method performs cross-validation on training sets using all possible parameter configurations within certain ranges. This method searches parameter space based on a "*try-all*" method, i.e. searching all the possible combinations and all ranges of different parameters specified by user for each kind of kernel. This way, the kernel and its parameter configuration with

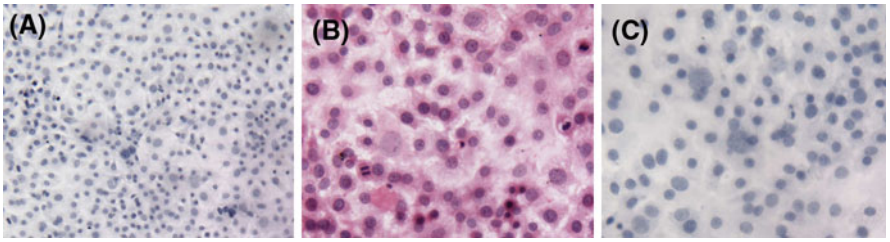

Fig. 3 NIH/3T3 fibroblasts. **a** NIH/3T3 fibroblasts culture stained with Haematoxylin and taken with a 10× objective (3t3-H-10×). **b** NIH/3T3 fibroblast cultures stained with Haematoxylin and Eosin and taken with a 20× objective (3t3-HE-10×). **c** NIH/3T3 fibroblast cultures stained with Haematoxylin and taken with a 20× objective (3t3-H-20×)

best performance (normally highest detection rate on training sets) is decided. In addition, the grid search method can also prevent over-fitting by finding the model with the best generalization [11].

We used the approach of Chang and Lin [3], to perform the grid search by providing two parameters to optimize a SVM classifier: the cost C and a kernel parameter γ [3,11]. The cost C is the penalty parameter on the error [2,4,21,24], a bigger Cost meaning giving a heavier penalty on errors. γ is a parameter used to configure kernels; for different kernel, γ has different meanings. In the RBF kernel, $exp(-\frac{||x-x_i||^2}{2\sigma^2})$, γ is $1/2\sigma^2$. A smaller γ will produce a more general classification boundary [4]. The optimal combination of these two parameters is found after a grid search which only needs be performed once for a given classification task.

3 Experiment

In this section, we explore the performance of SVM classifier on detecting cell nuclei from the background over a range of sample images.

3.1 Experimental material

The cells in the images were NIH/3T3 fibroblasts (a cell line) cultures grown on glass and stained with Haematoxylin (this stains the nuclei in blue and the cytoplasms are faintly stained also) or Haematoxylin and Eosin (blue and pink staining, the eosin stains the cytoplasms in pink). The images were taken with a QImaging Micropublisher 3.3 Firewire camera (providing a 2,048 × 1,536 pixel field attached to an Olympus BX50 microscope) with either a 10× objective (1 pixel = 1.239 μm) or a 20× objective (1 pixel = 0.624 μm). The images were background corrected and were the average of 8 image captures. We used both H and HE images. This was to investigate whether the performance would be affected by the use of deconvolution of two stains rather than one (HE stain method is more commonly used than H alone).

3.2 Training data

The data set consisted of 75 images. Among these, 25 images contained NIH/3T3 fibroblast cultures stained with Haematoxylin only (Fig. 3a). These 25 images were taken with a 10× objective (3t3-H-10×). 25 images were stained with Haematoxylin and Eosin (Fig. 3b) at 20× objective (3t3-HE-20×). The remaining 25 images also contained NIH/3T3 fibroblasts cultures stained with Haematoxylin (Fig. 3c), but they were taken with a 20× objective (3t3-H-20×). Three 3t3-H-10× fibroblasts cultures images were used to provide training samples. Each training image itself contains around 1500 examples of cell nuclei to use as individual training/testing examples.

A positive training set and a negative training set were created to train a SVM classifier. The cell nuclei become darkened after being stained. They are thus very easily manually extracted as positive samples from the background manually. For our cell detection task the positive training data are rectangular sub-images of part of individual 3t3 cells (the boundary of the sub-image is a rectangle, as it is impossible to include all the irregular shaped physical limits of the cytoplasms). So precisely speaking, our cell detection is a cell nucleus detection task which includes both cytoplasm and nucleus. Both positive and negative training samples in these training sets come from the three 3t3-H-10× training images. These samples were manually extracted from a subset of the original images (examples are shown in Fig. 4). Our rule of extracting positive samples was to include only one cell nucleus per sub-image, and align the centroid of each cell nucleus to the centre of the sub-image (e.g. Fig. 4). The width and height of each cell nucleus were no less than 2/3 of the width and height of the sub-image. Negative samples were manually extracted from image areas containing no cells (only background). Each sub-image may have different sizes according to the shape of its cell nucleus and were unified to a uniform size for training. All the extracted sub-images were cross-validated by four experience human experts participating in this research (authors). These initial training sets were further improved to form the final SVM classifier training sets. This is discussed further in Sect. 3.3.

3.3 Training the SVM classifier

An important issue with machine learning based detection is the selection of the input features used for detection that can distinguish targets (i.e. nuclei) from background. For our experiment, we mainly tested three types of inputs for a sub-image: raw pixel values, edge values, and the combination of these two. The edge value is calculated using the Laplace operator [23]. Other feature information considered were image moments, histograms [23] and their combination with raw pixel values but there were no clear improvements on the system performance compared with using only the raw pixel values or edge values. For the training, raw pixel values and Laplacian values were normalized to a range between 0 and 1. Their combination was simply the normalized sum of the two values. The reason for using edge information as an input component was that in a positive example the cell nucleus edge forms a circle and the edge strength is stronger than other parts of the image. In a negative example the edges distribute randomly. In order to isolate which input was the optimal choice, the grid search method was used to identify the input with the highest classification performance. In the grid search, we tested all available kernel types and their parameter configuration over certain ranges. The search range for Cost is from 0 to 10 with a step of 0.5. For Gamma is from 0 to 0.1 with a step of 0.005. The combination of raw pixel value and Laplacian value inputs gave the best performance with a RBF kernel selection. This combination improves the detection rate by at least 5% compared with other input values during the input selection procedure.

After the initial training sets (introduced in Sect. 3.2) were created, the next step was to train the SVM classifier. During the training, these sets would be improved continually by adding more key training samples to cover false-positive or false-negative instances. This improvement stops when the performance of the classifier stabilizes. The optimal kernel and its parameter configuration for the classifier could then be determined. Table 2 details the overall procedure of training the SVM classifier.

All input training samples were resized to a uniform image size for training the classifier. In order to improve training, detection speed and maintain high classification rate, 20×20 was empirically chosen as the uniform size. Figure 4 shows some samples of the training sub-images. Cell nuclei of positive training samples are located in the centres of the sub-images. Negative training samples are patches containing no nucleus, part of nucleus or more than one nucleus. If the centre of the nucleus is far away from the centre of the sub-image, this sub-image will also be treated as a negative sample. Table 3 shows quantity and size details of training data used in this work.

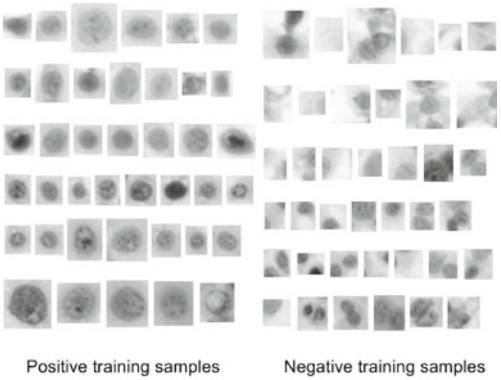

Fig. 4 Samples of sub-images for training a SVM classifier. *Left side* positive training samples. *Right side* negative training samples

Figure 5 shows the grid search for kernels and their optimized parameters in the final training iteration (of the procedure in Table 2) where the highest classifier performance has been reached. In this iteration, RBF kernel gives higher classification rate. It is taken for the SVM classifier. Alternative kernels, sigmoid and polynomial, were also included in the search but their performance was much lower than the RBF and linear kernels. The final results are discussed in more detail in Sect. 3.5.

3.4 Detecting cell nuclei

The trained SVM classifier was applied to detect cell nuclei from the unseen culture cell images. When searching for cell nuclei, a grid scan was performed across the image to extract image patches and return them to the SVM classifier for classification (thus returning both the classification result for cell nuclei presence at that sub-region patch position within the image). In order to cope up with nuclei of varying size, the size of the grid increased over multiple scan iterations.

Table 2 Procedure to train a SVM classifier

1. Use grid search method to determine appropriate kernel and optimize kernel parameters on initial training sets
2. Use selected kernel and its optimized parameters to train classifier using the initial training sets
3. Use trained classifier to detect cells from original images where the training samples are from
4. Update training sets by adding false positive sample to negative training set and false negative samples to positive training set
5. Use grid search method to determine appropriate kernel and optimize kernel parameters on updated training sets
6. Repeat step 3, 4, 5, until the detection is stable on those training images

Table 3 Training data details

Source images	Number of sub-images		Average sub-image size	Uniform size
	Positive	Negative	(width × height)	
3t3-H-10×	269	376	37 × 38	
3t3-HE-20×	232	311	32 × 34	20 × 20
3t3-H-20×	320	430	35 × 37	

Only the 3t3-H-10× is chosen for the final experiment. Other sets are only used for comparison purpose and they did not contribute to the performance of the classifier. (Training samples are resized to uniform size for training)

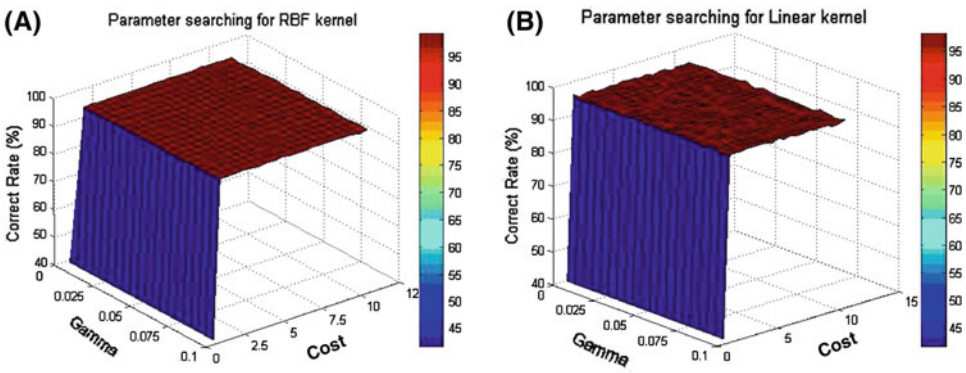

Fig. 5 Searching parameters for RBF and Linear kernels using grid search method. **a** RBF kernel. The highest classification rate is 99.1091% when Cost is 1.5 and Gamma is 0.08. **b** Linear kernel. The highest classification rate is 98.1069% when Cost is 2 and Gamma is 0.08 using the RBF kernel

Several iterations were performed based on defined starting and stopping cell nuclei scale criteria. Extracted image patches were rescaled to 20 × 20 pixels for SVM classification. The grid scan used an exhaustive pixel sliding window in the X and Y directions over the image defined by movement in N pixel steps in each direction. Post processing was performed to cluster neighbouring positive patches within a certain area and filter out smaller nuclei patches occurring inside large nuclei patches. Figure 6 illustrates the procedures for post-processing the detected cell nuclei. An important step is the tidy up of detections. To merge two overlapping rectangles, we calculate the overlapping area between them. If the overlapping area is bigger than 80% area of one of the rectangle, we count these two rectangles as a single detection and merge them together with a new rectangle which is the smallest one comprising those two rectangles. To remove a smaller rectangle within another one, we check if all four corners of the small rectangle are within the big rectangle. If they are, then this small rectangle will be discarded.

The system outputs are the coordinates and sizes of detected cell image patches. These outputs are post-processed values based on the sub-region patches that obtain a positive response from the SVM. They contain the cell nuclei. The user can do further calculations to get more precise cell locations, for example the geometrical centre of the cell, within these image patches by themselves.

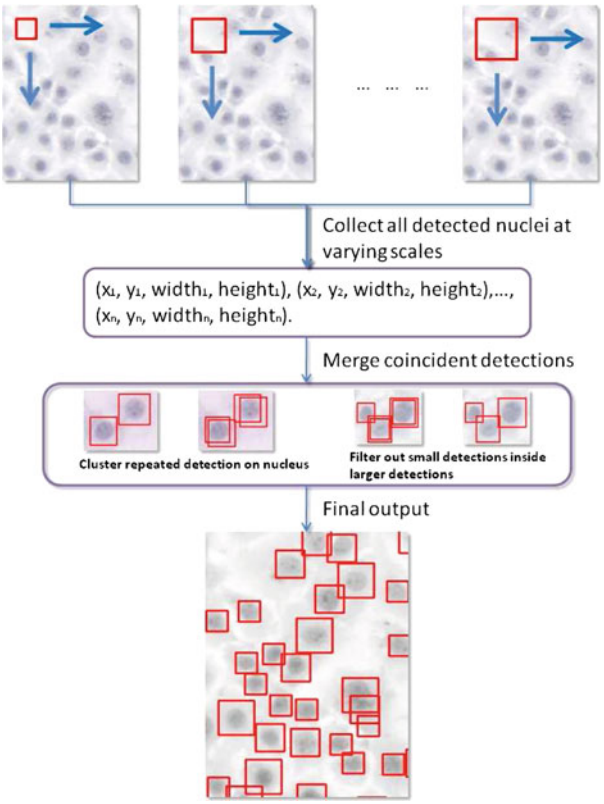

Fig. 6 The procedure for post-processing detected cell nuclei. The system finds cells with different sizes then consolidates them and forms final outputs

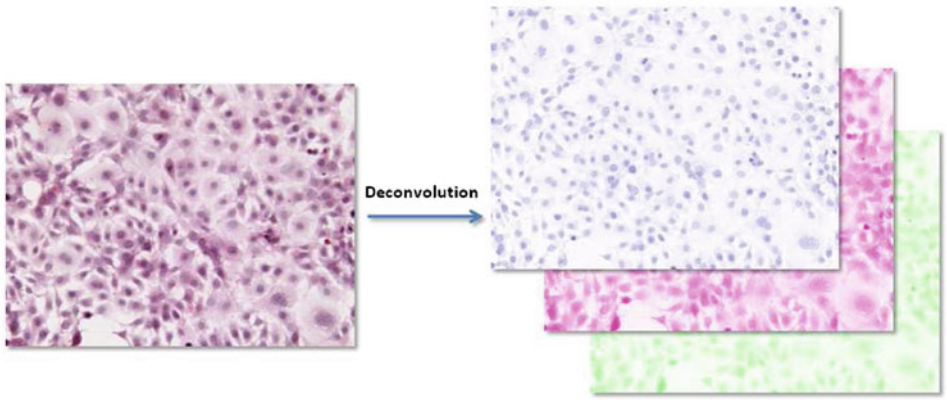

Fig. 7 Deconvolution algorithm extracts the Haematoxylin contribution from input images stained with Haematoxylin and Eosin. The Haematoxylin component of the image is shown on the top right

Table 4 Average detection results over the range of RBF-kernel SVM parameter settings illustrated in Fig. 5

Images	Detections	False negatives	False positives	Repeated detections	Accuracy (%)
3t3-H-10×	1410	41	3	66	92.24
3t3-HE-20×	432	13	2	22	91.15
3t3-H-20×	441	8	4	26	90.95

For the NIH/3T3 fibroblasts cultures stained with Haematoxylin and Eosin, we first applied H/E colour deconvolution [14,22] as a preprocessing step prior to the SVM classification. This algorithm deconvolves the colour information taken with RGB based cameras and splits the contribution of each stains based on stain specific RGB absorption [22]. In our experiment, we split the Haematoxylin contribution from input images stained with Haematoxylin and Eosin. This makes the input image colour information consistent with the training sub-images and removes other staining contribution to the input image which may influence the overall performance of the classifier. An example of this process is shown in Fig. 7.

3.5 Results

In the experiment, three 3t3-H-10× fibroblasts cultures images were used to provide training data and 72 images, made up of 22 3t3-H-10×, 25 3t3-H-20×, and 25 3t3-HE-20× fibroblasts cultures were then used to test the classifier detection performance. The results showed highly successful performance of the SVM classifier. In the experiment, we validated the classifier performance based on our own manual detection of cell nuclei. On these 2,048 × 1,536 images, the system achieved an average speed of 2 min/frame. The computer we used is a 2.66 GHz Core 2 Duo laptop with 3.45 GB of RAM.

During searching, we used different search parameters for 10× and 20× images, but constant parameters within each group. For the 10× images, the searching grid size started from 1.3×(training sample size) to 2×(training sample size). The increase step between scans at different scales was 0.3×(training sample size). For the 20× images, the searching grid size started from 2×(training sample size) to 3.5×(training sample size). The increase step between scans at different scales was 0.5×(training sample size). The detection intervals for both 10× and 20× images were 3 pixels on both X and Y directions. The average detection rate was above 90%. There were two 3t3-HE-20× images and two 3t3-H-10× images which were not well detected (80% detection). Adjusting the search parameters, the detection improved this to between 85 and 90%. If we continue tuning these parameters, the results would be further improved. The false positive detections were quite low. On average, it was less than 5 false positive detections in each image ($\sim$ 0.9% false positive per image based on $\sim$ 441 cells per 3t3-H-20 image). The set of average detection results over 72 test images is shown in Table 4. This table shows the average detections, false negatives, false positives and repeated detections within each image type (rounded to nearest integer). The results are classified according to image types.

Figures 8, 9 and 10 show samples of nuclei detection in 3t3-H-10×, 3t3-H-20×, and 3t3-HE-20× fibroblasts cultures images where we see the detection of nuclei over varying densities, scales, staining and types.

By means of comparison, we compared this result to the earlier machine learning based approach of [7] based on training over the same datasets as used in the work presented here and employing the methodology of [7] for solely cell nuclei detection. The resulting detection rate of the cascaded

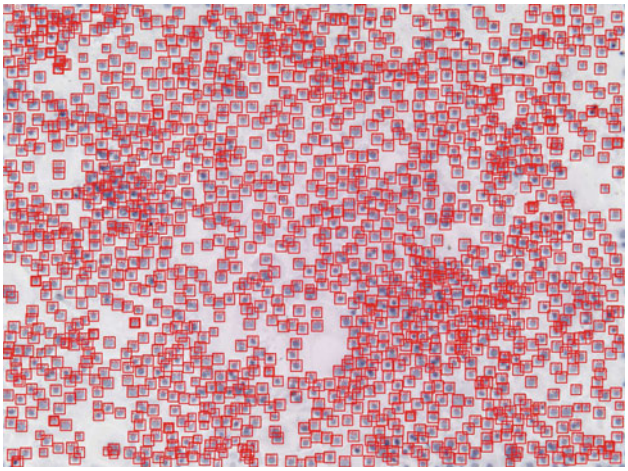

Fig. 8 Detection of 3t3-H-10× cell nuclei

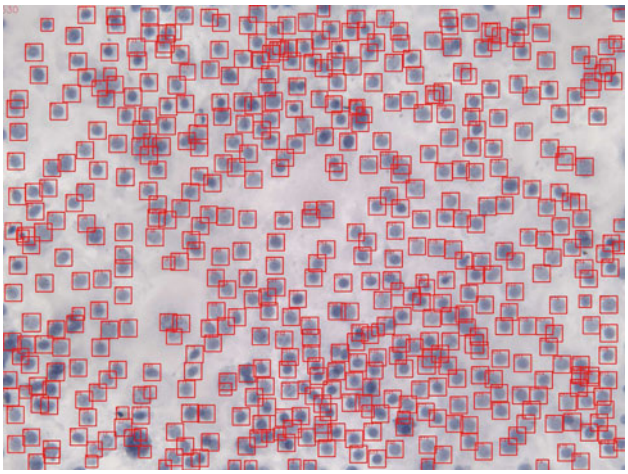

Fig. 9 Detection of 3t3-H-20× cell nuclei

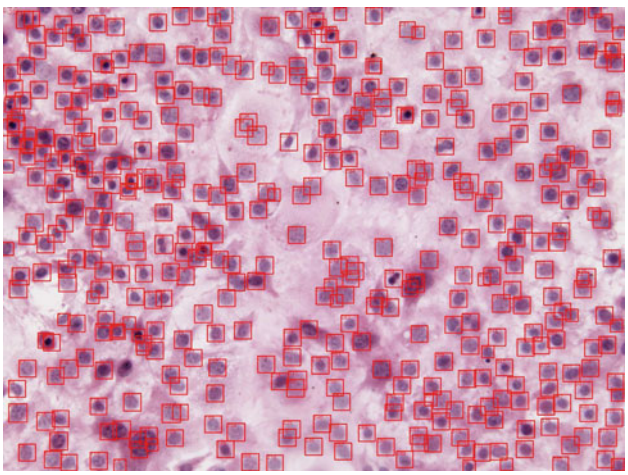

Fig. 10 Detection of 3t3-HE-20× cell nuclei

haar classifier approach of [7] is 82% which is considerably lower than that presented here using our novel Laplace-feature derived SVM methodology.

4 Discussion

The SVM classifier was able to successfully detect most of the individual 3t3 cell nuclei (∼ 90% success, Table 4). There were few false positive detections (≤0.9%, Table 4). Overall, the SVM algorithm displayed remarkable performance of generalization: we used only 1/25th of the available images and not all of the cell nuclei in these images as training data. We used only 3t3-H-10× images to train the classifier but it could detect 3t3-H-20× and 3t3-HE-20× nuclei well. Notably, the classifiers could discriminate individual nuclei even when they formed clusters or there were overlapping (see, Fig. 11a). In [7], we applied Haar like feature classifier to detect cyst using raw pixel values. Though it has comparable performance with other work [14], this classifier gave a higher false positive rate and missed some true positives (false negatives) when faced with complex cell clusters in comparison to the SVM approach proposed here. The performance of the SVM and the selection of input features reduced overall false detections (both positive and negative). In addition, using colour deconvolution to extract the Haematoxylin channel information did not seem to influence the nucleus detection performance in comparison to images of Haematoxylin-only stained cultures.

We noted that there were some repeated detection over some single cell nuclei. It is difficult for the classifier to detect nuclei successfully when the nucleus size is much larger than the class norm. This increased the repeated detection of an individual nucleus. These kind of false detections can be removed by providing a larger detection grid size range at the expense of and increased computational cost (Fig. 11b–d).

In addition to the kernel type and its parameters, the image search parameters also had significant influence on nucleus detection. These search parameters included the starting detection grid size, the grid size increasing factor and detection intervals. Currently, we have no formal method to selecting these parameters and chose them according to an empirically estimated minimum cell nucleus size. Figure 12 shows the influence of different search parameters. More cells with smaller size were detected using the 1.3 times grid size while a 1.4 times grid missed small cells. The differences have been marked using yellow lines. Circles 1, 2, and 4 show some cells detected with the 1.3 scale (45 cells are detected) but not the 1.4 scale (18 cells are detected) starting detection grid size.Circle 3 shows the different repeated detections. It is obvious the 1.3 starting detection search grid size gave better detection. Another interesting phenomenon was that the 1.3 times size gave less repeated detections than 1.4. This is probably related to the detection clustering algorithm, as it is more likely for a nucleus to be identified as being inside another and to be removed at the lower scale setting.

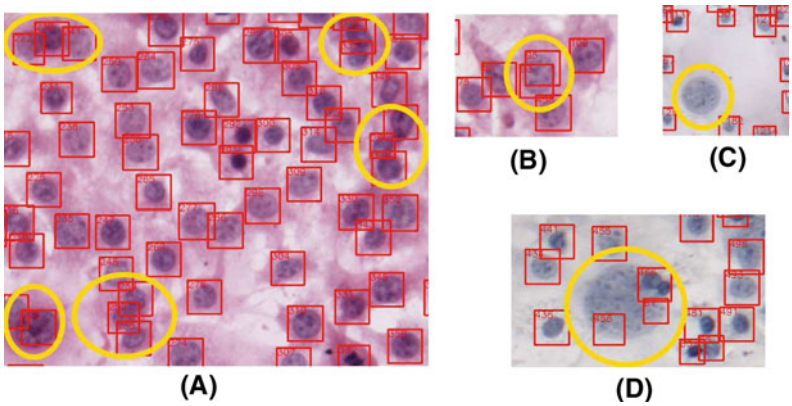

Fig. 11 **a** The classifiers could discriminate individual nuclei even when they formed tight clusters or there were overlaps. **b** A sample of repeated detections. **c** An undetected nucleus due to its large size. **d** A repeated cell nucleus due to its large size

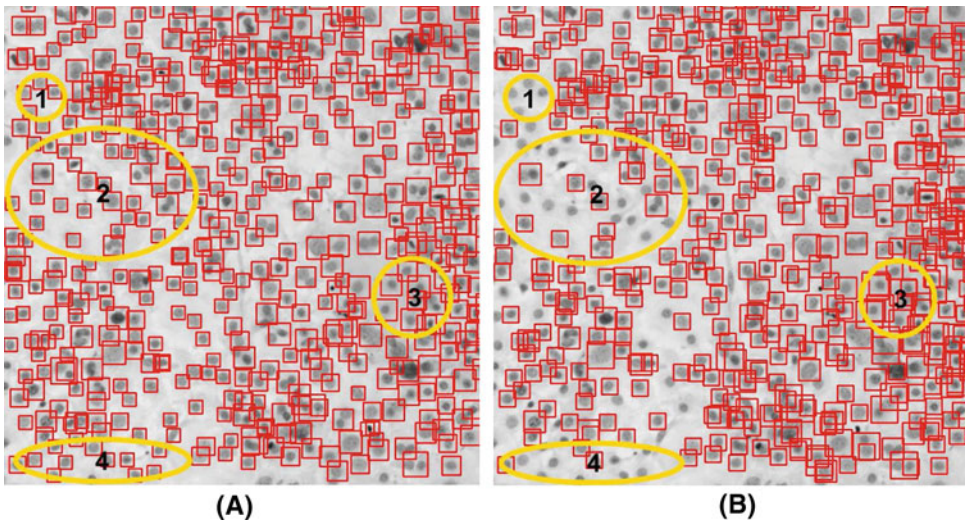

Fig. 12 Different starting detection grid sizes influence the cell detection. **a** A detection sample with starting detection grid size 1.3 times the training sample size. **b** is with 1.4 times the training sample size

5 Conclusions and future work

With the help of the grid search SVM training method, we can generate suitable, high performance classifiers, as the selection of classifier parameters has an important and clear influence on the system performance [11,16]. Input feature selection is crucial and here we combine both gradient edge and raw pixel values to provide a successful technique. Our approach shows a 90% average success rate based on an unpreprocessed raw image and generalizes over differences in scale and cell staining (with deconvolution used as appropriate). This supports the high performance of machine learning approaches [5,7,18] in this and other medical tissue classification problems. Notably the methodology outperforms earlier work in this field [7].

We are extending the current work to more complex cell images. Our current experiments used relatively regular images with small variation in cell shape and orientation. If the variation is large, the construction of training sets and the selection of features for detection is likely to be affected. Current training data is governed by the rectangular separation of cells from the original slide images. As such it is inevitable that the training samples include parts of other neighbouring cells which are likely to influence the training and possibly produce further false positive detections. In future, we will consider to use sub-parts of the nuclei as input features to locate possible cell nuclei with further verification to detect whole cell nuclei presence. A smarter search method will also be used though the speed now is acceptable. Currently, we are using a sliding window to search through the image for nuclei. Quick edge detection or threshold method may provide initial areas for search.

Acknowledgments This work is supported by the Science and Technology Facilities Council (ST/F003374/1 / ST/F003404/1), Cranfield University and the University of Birmingham.

References

1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. : Molecular Biology of the Cell. Garland Science, Oxford (2002)

2. Burges, C.J.C.: A tutorial on support vector machines for pattern. Data Mining Knowl. Discov. **2**, 121–167 (1998)
3. Chang, C.C., Lin, C.J.: Libsvm—a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm/
4. Cristianini, N., Shawe-Tayor, J.: An Introduction to Support Vector Machines. The Press Syndicate of The University of Cambridge, Cambridge (2000)
5. El-Naqa, I., Yang, Y.Y., Wernick, M.N., Galatsanos, N.P., Nishikawa, R.: Support vector machine learning for detection of microcalcifications in mammograms. In: IEEE Transactions on Medical Imaging, vol. 21, pp. 1552–1563 (2002)
6. Glotsos, D., Spyridonos, P., Petalas, P., Cavouras, D., Ravazoula, P., Dadioti, P., Lekka, I., Nikiforidis G.: Computer-based malignancy grading of astrocytomas employing a support vector machine classifier, the who grading system and the regular hematoxylin-eosin diagnostic staining procedure. Anal. Quant. Cytol. Histol. **26** (2004)
7. Han, J.W., Breckon, T., Randell, D., Landini, G.: Radicular cysts and odontogenic keratocysts epithelia classification using cascaded haar classifiers. In: McKenna, S., Hoey, J. (eds.) Proceedings of the Twelfth Annual Conference of Medical Image Understanding and Analysis, pp. 54–58. University of Dundee, UK (2008)
8. Han, J.W., Lane, P.C.R., Davey, N., Sun, Y.: Comparing the performance of single-layer and two-layer support vector machines on face detection. In: Proceedings of The Seventh UK Workshop on Computational Intelligence (2007)
9. Hearst, M.A.: Support vector machines. IEEE Intell. Syst. **13**, 18–28 (1998)
10. Herold, J., Friedenberger, M., Bode, M., Rajpoot, N., Schubert, W., Nattkemper, T.W.: Flexible synapse detection in fluorescence micrographs by modeling human expert grading. In: Proceedings of 2008 IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1347–1350. IEEE (2008)
11. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification. http://citeseer.ist.psu.edu/689242.html
12. Kohavi. R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, pp. 1137–1143 (1995)
13. Kruk, M., Osowski, S., Koktysz, R.: Recognition of colon cells using ensemble of classifiers. In: Proceedings of International Joint Conference on Neural Networks, pp. 288–293 (2007)
14. Landini, G.: Quantitative analysis of the epithelial lining architecture in radicular cysts and odontogenic keratocysts. Head Face Med. **2** (2006)
15. Landini, G., Othman, I.E.: Architectural analysis of oral cancer, dysplastic, and normal epithelia. Cytometry Part A **61 A**, 45–55 (2004)
16. Mitchell, T.M.: Machine Learning. The McGraw-Hill Companies, Inc (1997)
17. Mjolsness, E., DeDoste, D.: Machine learning for science: state of the art and future prospects. Science **293**, 2051–2055 (2001)
18. Nattkemper, T.W., Ritter, H.J., Schubert, W.: A neural classifier enabling high-throughput topological analysis of lymphocytes in tissue sections. IEEE Trans. Inf. Technol. Biomed. **5**, 138–149 (2001)
19. Nattkemper, T.W., Twellmann, T., Schubert, W., Ritter, H.: Human vs. machine: evaluation of fluorescence micrographs. Comput. Biol. Med. **33** (2003)
20. Rahman, Md.M., Desai, B.C., Bhattacharya, P.: Supervised machine learning based medical image annotation and retrieval. In: Accessing Multilingual Information Repositories. Lecture Notes in Computer Science, pp. 692–701. Springer, Berlin (2006)
21. Robinson, M., Castellano, C.G., Adams, R., Davey, N., Sun, Y.: Identifying binding sites in sequential genomic data. In: Artificial Neural Networks—ICANN 2007, pp. 100–109, Porto, Portugal. Springer, Berlin (2007)
22. Ruifrok, A.C., Johnston, D.A.: Quantification of histochemical staining by color deconvolution. Anal. Quant. Cytol. Histol. **23**, 291–299 (2001)
23. Russ, J.C.: The Image Processing Handbook. CRC Press, Inc, Boca Raton (1995)
24. Schölkopf, B., Smola, A.J.: Learning with Kernels. The MIT Press, Cambridge (2002)
25. Schnorrenberg, F., Pattichis, C., Kyriacou, K., Schizas, C.: Computer-aided detection of breast cancer nuclei. In: IEEE Trans. Inf. Technol. Biomed. **1**, 128–140 (1997)
26. Shear, M.: Cysts of the oral regions. 3rd edn. Wright, Oxford (1992)
27. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, p. 511, Los Alamitos, CA, USA, IEEE Computer Society (2001)
28. Wang, H., Zheng, C., Li, Y., Zhu, H., Yan, X.: Application of support vector machines to classification of blood cells. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi **20** (2003)
29. Wang, M., Zhou, X., Li, F., Huckins, J., King, R.W., Wong, S.T.C.: Novel cell segmentation and online svm for cell cycle phase identification in automated microscopy. Bioinformatics **24** (2008)
30. Wei, N., Flaschel, E., Friehs, K., Nattkemper, T.W.: A machine vision system for automated non-invasive assessment of cell viability via dark field microscopy, wavelet feature selection and classification. BMC Bioinform. **9**, 449 (2008)
31. Wei, N., You, J., Friehs, K., Flaschel, E., Nattkemper, T.W.: An in situ probe for on-line monitoring of cell density and viability on the basis of dark field micfroscopy in conjunction with image processing and supervised machine learning. Biotechnol. Bioeng. **97** (2007)