Radicular cysts and odontogenic keratocysts epithelia classification using cascaded Haar classifiers

Ji Wan Han[a*], Toby Breckon[a†], David Randell[b‡], and Gabriel Landini[b§]

[a]School of Engineering, Cranfield University, Cranfield, MK43 0AL, UK
[b]Oral Pathology Unit, School of Dentistry, The University of Birmingham,
St. Chad's Queensway, Birmingham, B4 6NN, UK

Abstract. The diagnosis of radicular cysts and odontogenic keratocysts is an important aspect of oral medical examination. In this paper, we investigate the novel use of cascaded Haar classifiers for this task. We present preliminary experimental results showing the success of these classifiers in locating individual cells/nuclei and in classifying the cyst sub-types. The results illustrate well the future potential for the application such data driven machine learning approaches into this arena.

1 Introduction

Cysts are pathological cavities filled with a fluid or semi-fluid contents and usually lined by epithelial tissue. Odontogenic cysts of the jaws are cysts in which the lining epithelium is derived from cells of the tooth forming tissues. With regards to the aetiology of these lesions, they are traditionally classified into two main groups: developmental (dentigerous, keratocysts, gingival cysts, etc.) and inflammatory (radicular, residual, paradental cysts). Radicular cysts are the commonest (in general arising from complications of dental caries), followed in frequency by dentigerous and odontogenic keratocysts (OKs) [1].

There are two important issues in relation to OKs. Firstly, they commonly show active epithelial growth and for this reason, some researchers believe that they should perhaps be regarded as neoplasms rather than cysts. Secondly, they are known to occur in two fashions: solitary (or sporadic) and as part of the Basal Cell Naevus (also known as Gorlin-Goltz's) Syndrome (BCNS). This syndrome is autosomal dominant with complete penetrance, variable expressivity and is characterised by the presence of multiple nevoid basal cell carcinomas of the skin, multiple Not necessarily synchronous) odontogenic keratocysts of the jaws, skeletal abnormalities, ectopic calcifications and plantar or palmar pits. The diagnosis of an OK is therefore an important event that should flag the need to further examination of other possible BCNS signs. It would be also useful to know whether the difference between the solitary and syndrome keratocysts can be achieved from their morphological/microscopical appearance, but this has remained an elusive problem in histopathology.

Landini [2] analysed epithelial lining architecture in radicular cysts and odontogenic keratocysts applying image processing algorithms to follow a traditional cell isolation based approach. This formed the basis for later estimation of tissue layer level and architectural analysis of oral epithelia [3, 4].

Recently, with the rapid development of machine learning algorithms, some medical image processing applications have adopted machine learning techniques to improve classification performance. Rahman *et al.* [5] applied Support Vector Machines to medical image annotation and retrieval. El-Naqa *et al.* [6] applied Support Vector Machines to detect microcalcifications in mammogram images, which outperformed all the other methods related.

In this paper, we will introduce a machine learning based approach (cascaded Haar classification [7]) to explore the feasibility of using such algorithms to classify *K*: solitary odontogenic keratocysts, *S*: basal cell naevus syndrome associated odontogenic keratocysts, and *R*: radicular cysts. In the preliminary results we present, this method shows successful detection of individual cell nuclei within the pathological slides in addition to promising classification rates on the cyst subtypes (*K*, *S* and *R*).

*j.w.han@cranfield.ac.uk
†toby.breckon@cranfield.ac.uk
‡d.a.randell@bham.ac.uk
§g.landini@bham.ac.uk

2 Using Haar cascade classifiers

2.1 Haar cascade classifier

The Haar cascade classifier was firstly proposed by Viola [7] and later improved by Lienhart [8] with a primary application domain for the detection of faces within the visual surveillance domain. The concept is to use a conjunctive set of weak classifiers to form a strong classifier - in this instance, a cascade of boosted classifiers applying Haar-like features (Figure 1). These Haar features are essentially drawn from the response of Haar basis functions to a given type of feature at a given orientation within the image (Figure 1). Individually, they are weak discriminative classifiers but when combined as a conjunctive cascade a powerful discriminative classifier can be constructed capable of recognising common structure over varying illumination, colour base and scale [7].

Figure 1. Haar-like features. A. Edge features. B. Line features. C. Centre-surround features.

This cascaded Haar classifier is trained using a set of multiple positive object (i.e. cell, face, car ...) and negative object image images (i.e. no cell, face, car etc.). The use of boosting techniques then facilitates classifier training to select a maximally discriminant subset of Haar features to act as a multi-stage cascade. In this way, the final cascaded Haar classifier consists of several simpler (weak) classifiers that all form a stage in the resultant complex (strong) classifier. These simpler classifiers are essentially decision-tree classifiers (with at least 2 leaves) that take the Haar-like feature responses as input to the weak classifiers and return a boolean 0,1 pass/fail response. A given region within the image must then achieve a pass response from all of the weak classifiers in the cascade to be successfully classified as an instance of the object the overall strong classifier has been trained upon. Figure 1 shows the Haar-like features used as the basis for the weak classifiers in this approach.

The Haar cascade classifier thus combines successively more complex classifiers in a cascade structure which eliminates negative regions as early as possible during detection but focuses attention on promising regions of the image. This detection strategy dramatically increases the speed of the detector [7, 8]. It has been successfully applied to solve some challenging computer vision tasks, such as real-time face detection [9] and pedestrian detection [10], both areas in which variations in illumination, scale and colour (clothing/race) are common problems for other techniques.

2.2 Training data

For our nuclear classification task the training data for *K*, *S*, and *R* cysts are images of lining epithelial cells, which include cytoplasm and nucleus (however the physical limits of the cytoplasms are in general not clear cut and so the limits of contiguous cells may not be precisely defined). The data was gathered from 150 colour digital images (x40 objective, pixel width=0.31 micrometres) of histopathology slides from the archives at the School of Dentistry at the University of Birmingham. The images (all containing epithelial cell lining) were from 9 cases of radicular cysts (*R*:45 images), 13 cases of solitary keratocysts (*K*:65 images) and 8 BCNS keratocysts (from 5 patients, *S*:40 images). These images were subdivided into individual cell/nuclear sub-images for training purposes (Figure 2) according to the location and size of the epithelial cells, using bounding rectangles. This was driven from information acquired using a prior watershed segmentation based method [2–4, 11]. Figure 2 shows the procedure of generating the sub-images for training from the prior watershed approach.

Most of these acquired sub-images could then be used as positive training samples, but due to factors like the variability in the intensity of the nuclear stain and the mismatch between the watershed segmented basins and the rectangular subimage frame, some of poor quality content images were manually removed. These unsuitable samples most frequently contained no contrasted nuclei or more than one nucleus inside the frame. Our rule of selecting positive samples was to include only one cell nucleus per sub-image. Other sub-types of architectural cell arrangements (i.e. more than 1 cell nucleus in the sub-images) will be dealt in future work. The negative training samples required were produced using randomly cut regions, of non-cell positions, from the whole cyst slides images. In addition positive training data from one cyst class can be utilised as negative training data for another (each of *K*, *S*, and *R*). An additional requirement of the algorithm was that all the training data images were normalised to a common size - in this preliminary study this was simply set as the mean size of the samples. Figure 3 shows some samples of accepted and rejected training

sub-images. Table 1 shows quantity and size details of training data used to train classifiers. When the classifiers are being trained, all training samples are resized to their average sizes shown in this table.

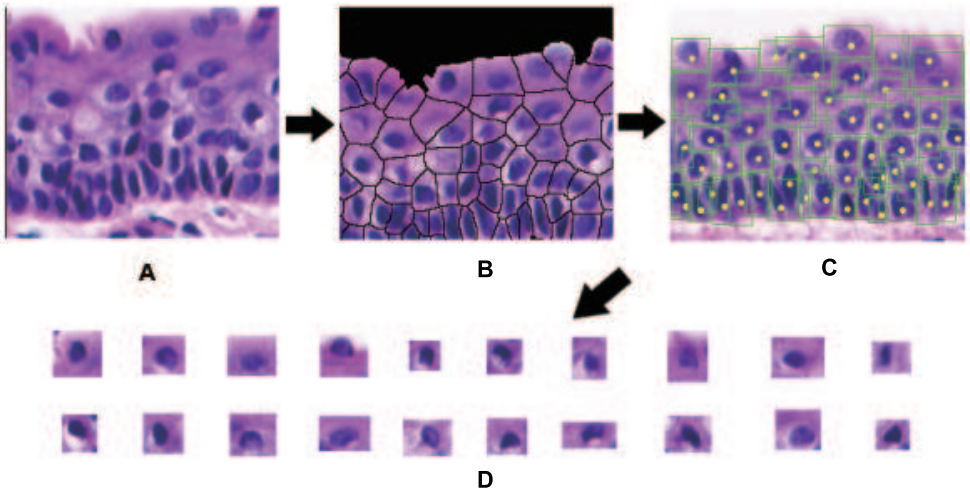

Figure 2. Generation of sub-images for training. A. Original cyst picture. B. Segmentation using watershed algorithm. C. Cell areas and their centre of mass (yellow dots). D. Extracted sub-images.

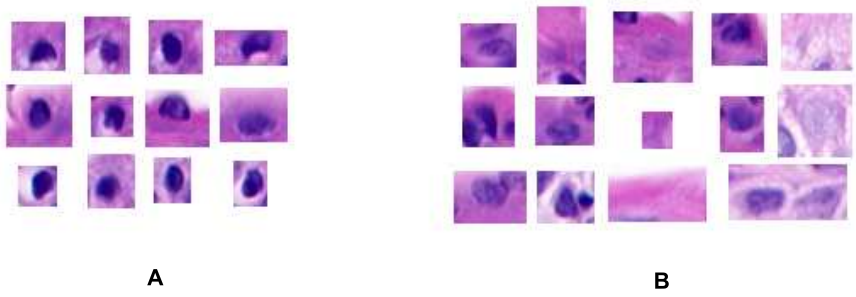

Figure 3. Selection of sub-images. A. Samples of accepted sub-images. There is normally only one nucleus in a sub-image. B. Samples of rejected sub-images. More than one nucleus or no nucleus in the sub image.

Cyst	Number of Sub-images		Average sub-image size
	Positive	Negative	$(Width \times Height)$
R	576	144	16×14
K	1254	144	14×15
S	478	144	15×15

Table 1. Training Data details. Training samples will be resized to the average sub-image sizes when classifiers are being trained.

2.3 Experimental method

We performed a primary exploration of the feasibility of using machine learning algorithms to classify Radicular cysts and Odontogenic Keratocysts based on the nuclear images. Using the previously detailed cascaded Haar classification approach three separate classifiers were created, respectively for *K*: solitary odontogenic keratocysts, *S*: basal cell naevus syndrome associated odontogenic keratocysts, *R*: radicular cysts. In operation all three classifiers then processed unseen histological images in turn to return a statistical count of the number of each corresponding cyst lining nuclei type present. From this, in a similar vein to manual human classification of such slides, we took the majority vote to classify the slide as containing pathological features corresponding to a given cyst type.

The training samples used were taken as sub-images from 18 *K*, 6 *R*, and 6 *S* cyst images. Separate images were then used for training validation and testing. More *K* cyst samples were used than the other two types of cysts because initially there were more false negative detections (cells could not be detected) than the other cyst types during the training and validation phase. We added more samples to improve the performance of *K* classifier to a level equal

to that of the other two classifiers - as a result - 19 *K*, 13 *R* and 11 *S* cyst images were used to test classifiers. The negative training samples from all three types of cysts were composed by background image patches without cells. The classifiers were trained using the data mentioned in 2.2. There were 14 stages in each classifier and the classifiers each used various types of Haar features including vertical and 45^o angular features [7, 8].

In the testing phase, we changed the classifier sizes to scan the image several times to find nuclei of different sizes by a factor of 1.1, which means the difference between two sequent classifier sizes is 10%. In addition, spatial clustering of co-incident nuclei detections was performed to eliminate repeated detection on the same nuclear region. A minimum number of 4 consistent neighbouring detections was required to confirm the presence of a given classification type.

2.4 Results

In the experiment, 30 cyst images were used to provide training data and 43 pictures were then used to test the classifiers based on the statistical majority of cell type present. The results showed to be promising at this preliminary stage. A total of 37 images were correctly classified as keratocysts (including *K* and *S* classes) or radicular cysts. Among them, 11 out of 13 *R*, 26 out of 30 *K* and *S* cysts were correctly classified giving a total correct classification rate of 86%. The classification rate between *K* and *S* cysts was 53.8% with 14 out of 26 were correctly classified. These results are comparable to those reported in [2], 100% correct classification rate between keratocyst and radicular cyst images but only 60% between *K* and *S* cysts, and seem to confirm that there are no characteristic morphological differences between the 2 OKs types. Figure 4 shows samples of nuclei detection in two samples of *K* cysts.

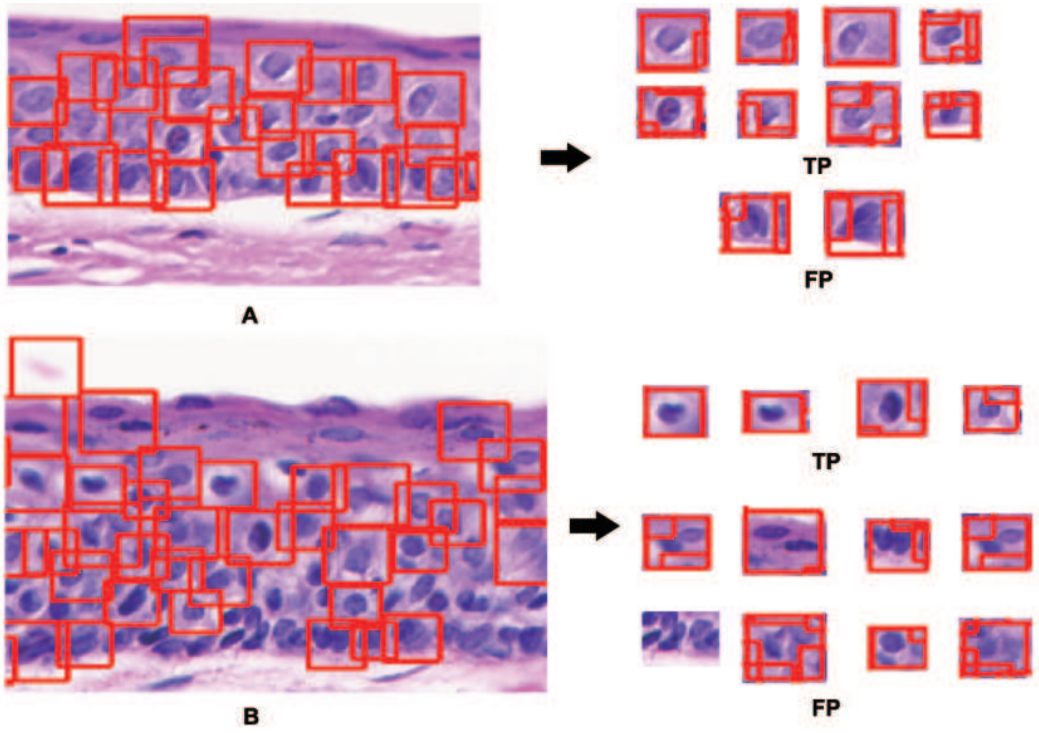

Figure 4. Detection of single nuclei. TP: True Positive. FP: False Positive. Both A and B are *K* cysts. The structure of B is more complex than that of A and in cyst A, detection of nuclei is relatively superior. More nuclei are missed in cyst B and more multiple nuclei per frame FP detections are present.

3 Discussion & Conclusions

The classifiers used were able to successfully find most of the individual epithelial cell nuclei, but there were many false positive detections. Notably, the classifiers could not discriminate individual nuclei when they formed tight clusters and therefore multiple nuclei false positives were common. In addition, we noted that it is difficult for the classifiers to detect nuclei successfully when the nuclear shape and orientation are extreme variants from the class norm. These false detections have a negative influence on the overall classification results of the technique. We hope that future work using an improved training data set (especially the negative examples) can help overcome this problem. Currently the methods of [2] consider tissue structural properties in addition to information on individual cells to give

a high overall correct classification result. However, the performance of this technique against that of [2] is based on lesser information. This is a promising result that guarantees further investigation and shows the potential for machine learning approaches in histopathology imaging problems.

4 Future work

Current training data is governed by the rectangular separation of cells from the original slide images. As such it is inevitable that the rectangles include parts of other neighbouring cells which are likely to influence the training and produce many false positive detections as a result (notably multiple nuclei false positives). This will be overcome by preprocessing the training data to remove such artefacts based on the *a priori* watershed segmentation of the training images instead of using rectangular sub-images. Secondly, we intend to introduce a secondary constraint related to the distance between cells. So far, we only used the individual cell/nuclear information without considering tissue structure information [4]. Our future work will combine both of these aspects to improve overall classification performance of the technique.

Acknowledgements

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. M. Shear. *Cysts of the oral regions*. Wright, Oxford, third edition, 1992.
2. G. Landini. "Quantitative analysis of the epithelial lining architecture in radicular cysts and odontogenic keratocysts." *Head & Face Medicine* **2**, 2006.
3. G. Landini & I. E. Othman. "Architectural analysis of oral cancer, dysplastic, and normal epithelia." *Cytometry Part A* **61A**, pp. 45–55, 2004.
4. G. Landini & I. E. Othman. "Estimation of tissue layer level by sequential morphological reconstruction." *Journal of Microscopy* **209**, pp. 118–125, 2003.
5. M. M. Rahman, B. C. Desai & P. Bhattacharya. "Supervised machine learning based medical image annotation and retrieval." In *Accessing Multilingual Information Repositories, Lecture Notes in Computer Science*, pp. 692–701. Springer Berlin/Heidelberg, 2006.
6. I. El-Naqa, Y. Y. Yang, M. N. Wernick et al. "Support vector machine learning for detection of microcalcifications in mammograms." In *IEEE Transactions on Medical Imaging*, volume 21, pp. 1552–1563. 2002.
7. P. Viola & M. Jones. "Rapid object detection using a boosted cascade of simple features." In *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, p. 511. IEEE Computer Society, Los Alamitos, CA, USA, 2001.
8. R. Lienhart & J. Maydt. "An extended set of haar-like features for rapid object detection." In *2002 International Conference on Image Processing. 2002. Proceedings.*, volume 1, pp. 900–903. 2002.
9. P. Viola & M. J. Jones. "Robust real-time face detection." *International Journal of Computer Vision* **57**, pp. 137–154, 2004.
10. P. Viola, M. J. Jones & D. Snow. "Detecting pedestrains using patterns of motion and appearance." *International Journal of Computer Vision* **63**, pp. 153–161, 2005.
11. L. Vincent & P. Soille. "Watersheds in digital spaces: an efficient algorithm based onimmersion simulations." In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 13, pp. 583–598. 1991.