# ADAPTIVE CELL DETECTORS FOR HISTOPATHOLOGICAL IMAGE ANALYSIS

*Bin Kong, Annika Fredrikson, Xiaofan Zhang and Shaoting Zhang*
Department of Computer Science, University of North Carolina at Charlotte, NC, USA

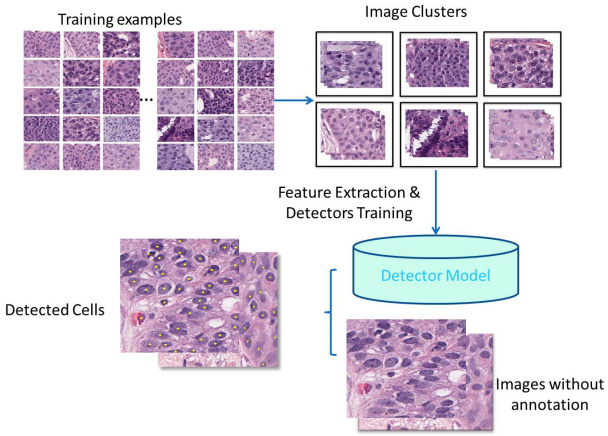


**Fig. 1**. Overview of our proposed framework.

**Motivation:** Breast cancer imposes a major threat on women's health. Fortunately, early diagnosis and treatment can significantly increase the survival rate among patients. In this context, computer-aided diagnosis (CAD) systems have been widely used in an attempt to relieve the workload on pathologists and to offer more reliable and consistent analysis of biopsy images [1]. To conduct such automated analysis, it is necessary to detect cells in these images. Recently, Arteta, et al. [2] proposed to segment possible cell nuclei using maximally stable extremal regions (MSER) images and structure support vector machine (SVM) to classify the detection candidates. Despite its efficacy with Phase-contrast images, we observed that above framework does not work well when coping with histopathological images acquired from breast microscopic tissues. The reason is that these images are heterogeneous and greatly differ from each other. In this work, we improve the above mentioned Oxford-detector [2] with a clustering scheme for robust and adaptive cell detection.

**Methodology:** Fig.1 shows our presented framework of adaptive detector framework. Because of the large variance of appearance among these images, simply training classifiers using all images can lead to unsatisfied result. In our solution, we extract high-dimensional textural features of the images based on bag-of-colors. Using these effective features, we can group the training images into several image clusters. Each image cluster contains certain type of images which have similar appearance. Then, we train a specific detector on each image cluster using soft linear SVMs trained with liblinear because of its efficiency and accuracy. During runtime cell detecting, bag-of-colors features are extracted from the detecting images and these features are used to locate the most similar cluster, where the corresponding detector is used.

*Image Clustering:* Given a set of training images $\mathcal{I}^1$,$\mathcal{I}^2$, $\ldots\mathcal{I}^M$, where each training image $\mathcal{I}^j$ generates a high-dimensional textural feature $\mathbf{f}^j$ using bag-of-colors. Based on these effective and discriminative features, we can employ k-means or affinity propagation algorithm to generate $N$ clusters. *Adaptive Detectors:* Histopathological images are clustered into different groups. Suppose we have $M$ bag-of-colors features $\mathbf{f}^1$,$\mathbf{f}^2$, $\ldots\mathbf{f}^M$ which are extracted from $M$ training images, and we cluster them into $N$ clusters. The cluster centroid is denoted by $\boldsymbol{\mu}_1$,$\boldsymbol{\mu}_2$,$\ldots\boldsymbol{\mu}_N$. Given a new image, its bag-of-colors feature is extracted and corresponding cluster center is located by matching with the cluster centroid. The detector corresponding to that cluster is employed for adaptive cell detection.

**Experiments:** 300 histopathological images (around 1.6K pixels per image) of breast tissues are used as training examples, including usual ductal hyperplasia (UDH), atypical ductal hyperplasia (ADH), and ductal carcinoma in situ (DCIS). 250 color words are generated from each training exampleThese features are clustered into 6 categories, each of which is used to train a soft linear SVM. 20 images are employed to test our framework. The Oxford-detector is compared as the baseline. The precision and recall rates demonstrate that our framework significantly improves the detection accuracy (around 80% v.s. 60%), which is consistent with our motivation to train different classifiers based on different image clusters. In the future, we will create additional clusters to see whether it improves the results or not. Moreover, we will incorporate other features, such as Histogram of Oriented Gradients (HOG), into the clustering process to improve the clustering effects. We will also incorporate the detection results into our large-scale image retrieval framework [3]. The majority of this work was supported by the National Science Foundation Award #1156822, as part of a Research Experience for Undergraduates (REU) Site, hosted at the University of North Carolina at Charlotte and the Video and Image Analysis Lab.

## 1. REFERENCES

[1] Metin N Gurcan, Laura E Boucheron, Ali Can, Anant Madabhushi, Nasir M Rajpoot, and Bulent Yener, "Histopathological image analysis: A review," *TBME*, vol. 2, pp. 147–171, 2009.

[2] Carlos Arteta, Victor Lempitsky, J Alison Noble, and Andrew Zisserman, "Learning to detect cells using non-overlapping extremal regions," in *MICCAI*, pp. 348–356. Springer, 2012.

[3] X Zhang, W Liu, M Dundar, S Badve, and S Zhang, "Towards large-scale histopathological image analysis: Hashing-based image retrieval," *TMI*, vol. 34, pp. 496–506, 2014.