Tissue Classification using Cluster Features for Lesion Detection in Digital Cervigrams

Xiaolei Huang[1], Wei Wang[1], Zhiyun Xue[2], Sameer Antani[2], L. Rodney Long[2], Jose Jeronimo[3]

[1] Department of Computer Science and Engineering, Lehigh University, PA, USA

[2] Communications Engineering Branch, National Library of Medicine, MD, USA

[3] Division of Cancer Epidemiology and Genetics, National Cancer Institute, MD, USA

ABSTRACT

In this paper, we propose a new method for automated detection and segmentation of different tissue types in digitized uterine cervix images using mean-shift clustering and support vector machines (SVM) classification on cluster features. We specifically target the segmentation of precancerous lesions in a NCI/NLM archive of 60,000 cervigrams. Due to large variations in image appearance in the archive, color and texture features of a tissue type in one image often overlap with that of a different tissue type in another image. This makes reliable tissue segmentation in a large number of images a very challenging problem. In this paper, we propose the use of powerful machine learning techniques such as Support Vector Machines (SVM) to learn, from a database with ground truth annotations, critical visual signs that correlate with important tissue types and to use the learned classifier for tissue segmentation in unseen images. In our experiments, SVM performs better than un-supervised methods such as Gaussian Mixture clustering, but it does not scale very well to large training sets and does not always guarantee improved performance given more training data. To address this problem, we combine SVM and clustering so that the features we extracted for classification are features of clusters returned by the mean-shift clustering algorithm. Compared to classification using individual pixel features, classification by cluster features greatly reduces the dimensionality of the problem, thus it is more efficient while producing results with comparable accuracy.

Keywords: Image segmentation, color space, support vector machines, clustering, features, tissue classification, lesion detection, digital cervigrams, cervical cancer

1. INTRODUCTION

To make images searchable by content in large medical archives, it is very important to reliably segment and label different tissue regions, especially biomarker regions. We consider the automated segmentation problem in a very large archive of 60,000 digitized uterine cervix images, created by the National Library of Medicine (NLM) and the National Cancer Institute (NCI). These images are optical cervigram images acquired by Cervicography using specially designed cameras for visual screening of the cervix, and they were collected from the NCI Guanacaste project for the study of visual features correlated to the development of precancerous lesions. The most important observation in a cervigram image is the Acetowhite (AW) region, which is caused by whitening of potentially malignant regions of the cervix epithelium, following application of acetic acid to the cervix surface.

The cervigram images in the archive have large variations in their appearance due to illumination variations, artifacts in image acquisition, and intrinsic differences in image content. Existing methods for cervigram image analysis [1,2,3] consist of sequential steps of image processing such as pre-processing to remove specular reflection, segmenting
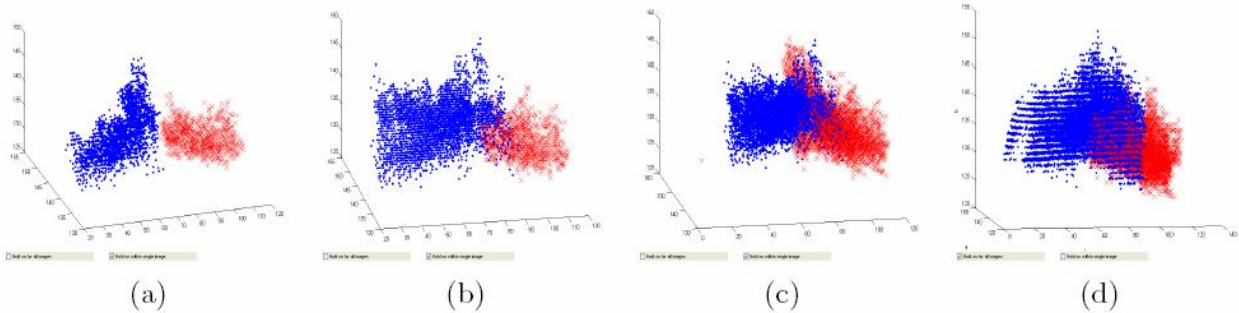
Figure 1. AW vs. Non-AW Cervix color sample distributions in L*a*b* space. (a) samples from one image. (b) samples from two images. (c) samples from three images. (d) samples from six images. (red) AW color samples. (blue) Cervix color samples.

cervix boundary, detection of OS, detection of columnar epithelium, thresholding, and mosaic texture analysis. These methods achieved good results on cervix area detection and on filling in specular regions, but the performance on segmentation of important tissue regions such as acetowhite and columnar epithelium needs improvement. Furthermore, due to large variations in image appearance, color and texture features of a tissue type in one image often overlap with that of a different tissue type in another image. This makes reliable segmentation in a large number of images a very challenging problem.

A popular color-based tissue segmentation method is to apply clustering techniques such as K-means [3], Gaussian Mixture Models [2] and Mean-shift [6], to directly model the posterior probabilities $p(c|e)$, where e represents an evidence vector that describes image features (e.g. pixel color), and $c = 1, ...,C$ is one of the C tissue classes. One challenge facing clustering methods in large-scale segmentation is that color distribution of one tissue class from many images can have many modes and overlap significantly with color distributions of other tissue classes. Figure 1 demonstrates this problem by displaying Acetowhite (AW) and cervix (non-AW) color samples from 1, 2, 3, and 6 images. Note that, as the number of images increases, the AW and cervix samples increasingly overlap with each other. It is therefore difficult to predict the class of a test color sample without a high probability of error given assumptions about the color distributions of the tissue classes.   In addition, not every tissue type is always present in every image, hence there will lack reliable ways to automatically set clustering parameters such as the number of clusters in K-means and GMM, and the size of the bandwidth in Mean shift.

In this paper, we propose a database-guided segmentation paradigm in which we apply machine learning techniques, such as support vector machines (SVM) to learn, from a database with ground truth annotations provided by experts, critical visual signs that correlate with important tissue types and to use the learned classifier for tissue segmentation in unseen images. The support vector machines (SVM) classifier [4, 5] has been successfully applied to detecting Microcalcifications in Mammograms [8] and various other medical classification problems. In this paper we use SVM to perform color-based tissue classification in order to segment different tissue regions, especially to segment the biomarker acetowhite (AW) region from the rest of the cervix.   The segmentation performance is optimized with respect to the feature color space and granularity. We evaluated color spaces including RGB, HSV, and L*a*b*. On different granularity of the features, we train AW and other tissue classifiers, first using individual pixel sample colors and then using cluster features returned by the Mean Shift based clustering algorithm [6].   Cluster features greatly reduce the dimensionality of training so that SVM is scalable to larger training sets, while producing results with comparable accuracy. Given a novel test image, the Mean Shift clustering algorithm partitions the image into clusters of similar color and/or texture, and the trained SVM classifier (on cluster features of training data) is applied to classifying clusters in the test image.   This ground-truth database guided segmentation method is flexible in terms of the number of tissue classes. Thus we can perform either two-label (e.g. AW vs. Non-AW cervix), or multi-label (e.g. AW, CE, SE, other) classification.

2. METHODOLOGY

For training and testing purposes, we have access to ground truth boundary markings on 939 cervigram images from the NCI/NLM archive. The ground truth markings are collected using a web-based Boundary Marking Tool developed by NLM and NCI [7]. There were 20 expert evaluators who used the tool to manually outline AW and Cervix boundaries in the 939 cervigrams. Some cervigrams have boundaries annotated by multiple experts; in this case, we randomly choose one as the ground truth although mechanisms for combining multiple-expert annotations are available [9]. Fig. 1 shows some example cervigram images with expert annotations.

Figure 2. Example cervigram images with boundary markings by experts.   The blue outlines AW regions and the yellow outlines cervix boundary.
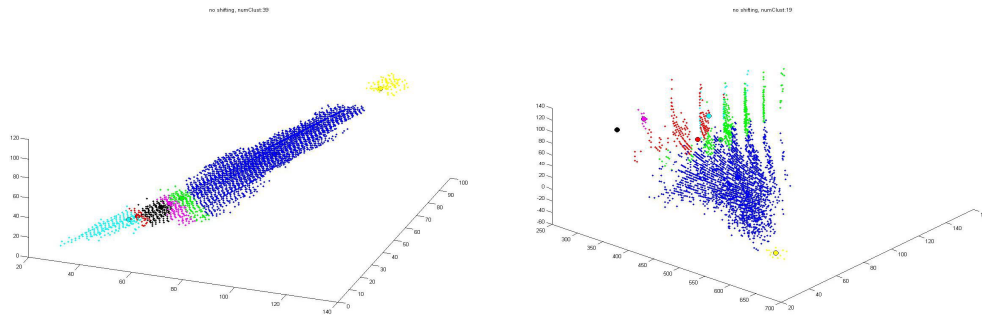

Figure 3.   Distribution of pixel color samples in RGB (left) and L*a*b* (right) color spaces.

2.1. Color Features

Color and texture are two most prominent features for tissue classification in digital cervigrams. In this paper, we investigate the color features in different color spaces and at different granularities.

Choices of color spaces include the RGB, HSV, CIE L*a*b*, and others. Our experiments show that luminance (or intensity) is an important feature in discriminating between tissue classes, thus HSV and L*a*b* are preferred color spaces.   Because of the quantization discontinuity in the hue dimension (e.g. 255 and 0 hue values are both very close to red) in the HSV color space however, for segmentation we choose to use pixel colors in the CIE L*a*b* space. Distribution of color samples in the L*a*b* space is also better for clustering and classification [6], as shown in Fig. 2.

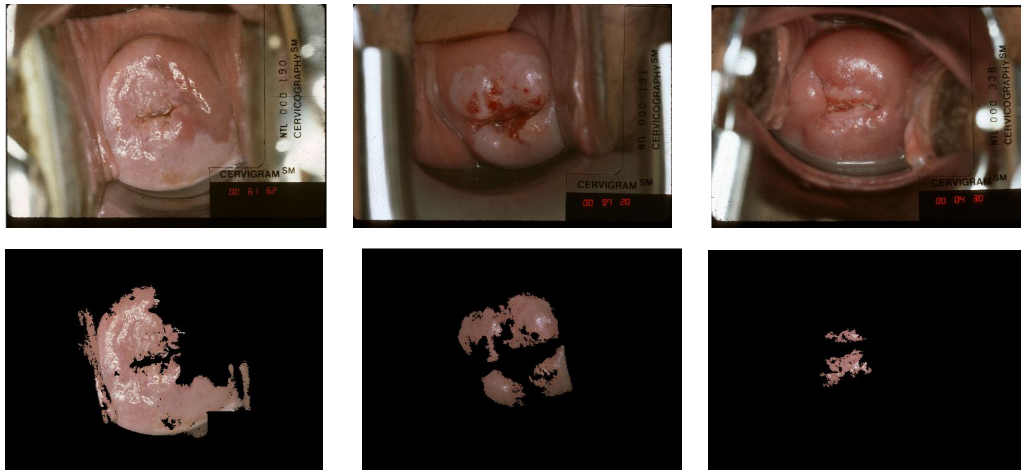**Figure 4.** AW segmentation examples. Test image (top row), segmented AW regions (bottom row)

SamplesPerImage * # of Images	Training Time	Memory Usage
200*10	67 sec	16MB
200*20	18 mins	53MB
200*30	1 hour 34 mins	103MB
200*55	20 hours 20mins	362MB

Table 1. SVM training time and memory usage given training sets of different sizes.

2.2. Pixel sample features vs. Cluster features

2.2.1 Pixel color features for training and classification

A two-class pixel-wise color classifier is trained by selecting Acetowhite (AW) pixel colors in the marked ground truth AW areas as positive samples, and selecting Cervix (non-AW) pixel colors as negative samples. 55 images are used for training and 120 images for testing. A confidence-rated SVM classifier with a linear kernel [4] is trained to differentiate AW from non-AW pixel colors. Given a test image, the classifier is applied directly to image pixels, and all pixels having the confidence rate above zero are considered as part of the AW region. Using false positive fraction (FPF) and false negative fraction (FNF) for quantitative evaluation, the pixel-wise classifier achieved an average of 23% FPF and 9% FNF. Some examples of AW segmentation using the pixel-wise classifier are shown in Fig. 4.

2.2.2 Mean-shift clustering and using cluster centers for training and classification

The SVM learning on pixel color features produces promising AW classification results as shown above. However, before SVM learners can potentially become a solution for tissue (especially biomarker tissue) segmentation in large medical image archives, we need to address the scaling problem. We record in Table 1 the processing time and memory usage given training sets of different sizes. One can see that, as the number of images and/or the number of pixels from each image for training increase, the SVM training time and memory usage explode exponentially.

To solve this problem, we experiment with cluster or region features instead of individual pixel features. For each training image, we first apply mean-shift clustering based on L*a*b* color feature and spatial proximity [6] to group pixels in the image into clusters. The label of each cluster (AW or Cervix non-AW) is assigned automatically based on
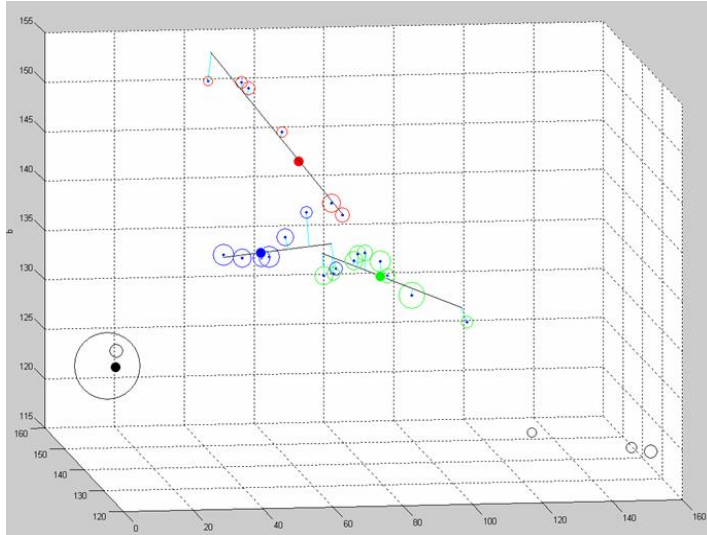
Figure 5. Labeled primary cluster centers of a training image.

Feature	SamplesPerImage * # of Images	Preprocessing (clustering) time	SVM Training Time	Memory use
Pixel color	200*55	0	20 hours 20mins	362MB
Cluster mean color	30*55	36 mins	< 1 min	12MB

Table 2. Performance comparison: cluster centers as training data vs. pixel colors as training data

expert boundary markings. The cluster center, which refers to the mean color of all pixels in the cluster, is then taken as the training sample. Fig. 5 shows labeled cluster centers for different tissues in one training image. Using approximately 30 cluster centers (from 15 largest AW clusters and 15 largest Cervix clusters) per image, the training time and memory cost are significantly reduced (Table 2).

Given a test image, it is first partitioned into clusters using mean shift. Then the SVM classifier learned on cluster centers is applied to classifying each cluster in the test image (using the cluster center feature) to either AW or Cervix. In our experiments, the segmentation accuracy on AW using cluster-based classification is comparable to that using pixel-based classification, while cluster classification is much more efficient and requires less memory (see Tables 1 and 2).

2.3. Multiple-label Classification

Instead of two classes (AW vs. cervix), we perform multi-label classification to segment simultaneously several significant tissue regions in cervigrams including the Acetowhite (AW), Columnar Epithelium (CE) and Squamous Epithelium (SE). The multi-label classifier is learned based on the "one-against-one" approach [10]. First, k*(k-1)/2 classifiers are trained, each using data from two different classes. In a voting strategy, each binary classification by a two-class classifier is considered to be a voting where votes can be cast for all data points. In the end, a data point is

Figure 6. Test result for multiple tissues by using RBF kernels of SVM. First is the original image; second is AW; third is CE.
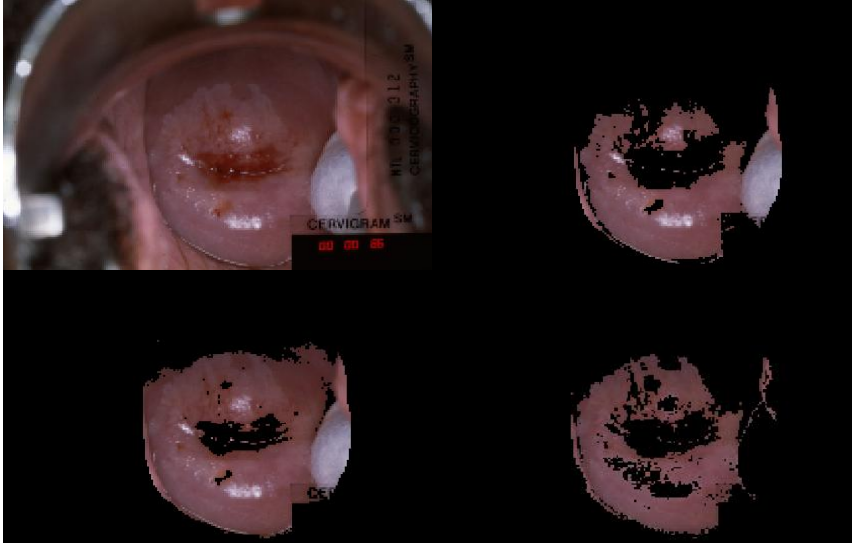

Figure 7. Using different kernels of SVM for two-label (AW vs. cervix) classification. Original image (upper left), AW by linear kernel (upper right), AW by polynomial with d=3 (lower left), RBF kernel (lower right).

labeled to be in the class with maximum number of votes. One example output of segmented AW and CE regions is shown in Fig. 6.

2.4. Kernel function selection of SVM

We consider several kernel function selections for the support vector machines classifiers.

Linear:
$$K(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{x_i}^T \mathbf{x_j}.$$

Polynomial:
$$K(\mathbf{x_i}, \mathbf{x_j}) = (\gamma \mathbf{x_i}^T \mathbf{x_j} + r), \gamma > 0.$$

Radial basis function (RBF):
$$K(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \| \mathbf{x_i} - \mathbf{x_j} \|^2), \gamma > 0.$$

Sigmoid:
$$K(\mathbf{x_i}, \mathbf{x_j}) = \tanh(\gamma \mathbf{x_i}^T \mathbf{x_j} + r).$$

Our empirical studies show that SVM classification is sensitive to kernel selection, especially when we use cluster-center features because the number of samples for training is fewer. Fig. 7 shows an example demonstrating result differences by different kernels. Using cluster center features, our experiments show that the RBF kernel outperforms others in two-label classification, while the linear kernel is the best in multi-label classification. Using individual pixel color features, the segmentation accuracy on most tissues, such as AW, CE, and SE, is comparable by different kernels on most test images.

3. CONCLUSIONS AND DISCUSSIONS

We introduce a database guided discriminative approach to segmenting tissue, especially biomarker acetowhite tissue, regions in digitized uterine cervix images. Training a support vector machine (SVM) classifier using cluster center features gives us better efficiency than using individual pixel features due to the reduced dimensionality while producing comparable accuracy. The method can be extended to segmenting other significant tissue regions in cervigrams including the Columnar Epithelium (CE) and Squamous Epithelium (SE) using multiple label classification. Comparing different kernel functions for the support vector machines classifier, we find that, with cluster features, the linear kernel is more suitable in multi-label classification while the Radial Basis Functions kernel is better for two-label classification.

ACKNOWLEDGMENTS

REFERENCES

[1] Zimmerman G. and Greenspan, H., "Automatic detection of specular reflections in uterine cervix images," Proc. of SPIE Medical Imaging, volume 6144, pages 2037- 2045 (2006)

[2] Gordon S., Zimmerman, G., Long, R., Antani, S., Jeronimo J. and Greenspan, H., "Content Analysis of Uterine Cervix Images: Initial steps towards content based indexing and retrieval of cervigrams," Proc. of SPIE medical imaging, volume 6144, pages 1549-1556 (2006)

[3] Tulpule, B., Hernes, D., Srinivasan, Y., Yang, S., Mitra, S., Sriraja, Y., Nutter, B., Phillips, B., Long, L.R. and Ferris, D., "A probabilistic approach to segmentation and classification of neoplasia in uterine cervix images using color and geometric features," Proc. of SPIE Medical Imaging, Volume 5747, pages 995–1003 (2005)

[4] Joachims, T., "Making large-Scale SVM Learning Practical," Advances in Kernel Methods - Support Vector Learning, B. Schölkopf and C. Burges and A. Smola (ed.), MIT Press (1999)

[5] Vapnik, V., "Statistical Learning Theory," Wiley-Interscience (1998)

[6] Comaniciu, D. and Meer, P., "Mean shift: A robust approach toward feature space analysis," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5 (2002)

[7] Jeronimo, J., Long, L., Neve, L., Bopf, M., Antani, S. and Schiffman, M., "Digital tools for collecting data from cervigrams for research and training in colposcopy," J. of Lower Genital Tract Disease 10(1) (2006) 16–25.

[8] El-Naqa, I., Yang, Y., Wernick, M. N., Galatsanos, N. P. and Nishikawa, R. M., "A Support Vector Machine Approach for Detection of Microcalcifications," IEEE Trans. on Medical Imaging, 21(12): 1552-1563 (2002)

[9] Warfield, S. K., Zou, K. H. and Wells, W. M., "Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation," IEEE Trans. on Medical Imaging, 23(7): 903-921 (2004)

[10] Knerr, S., Personnaz, L. and Dreyfus, G., "Single-layer learning revisited: a stepwise procedure for building and training a neural network," In J. Fogelman, editor, Neu-rocomputing: Algorithms, Architectures and Applications. Springer-Verlag (1990)