

# Automated Classification of Human Histological Images, A Multiple-Instance Learning Approach

Dehua Zhao<sup>1</sup>, Yixin Chen<sup>1\*</sup>, and Hernan Correa<sup>2</sup>

<sup>1</sup>Department of Computer Science, University of New Orleans

<sup>2</sup>Department of Pathology, Louisiana State University Health Sciences Center

<sup>1</sup>{dzhao, yixin}@cs.uno.edu, <sup>2</sup>hcorre@lsuhsc.edu

**Abstract**—In this paper, we apply a multiple-instance learning (MIL) method, MILES (Multiple-Instance Learning via Embedded instance Selection), to human histological image classification. MILES converts a MIL problem to a supervised learning problem by an instance-based feature mapping. 1-norm SVM is then adopted to select features and construct a classifier simultaneously. MILES identifies the sub-images that reflect underlying category concepts, and use them for classification. Experimental validation is provided based on images from different organs and parts of the body. The new approach demonstrates significantly improved performance in comparison with a method based on a Gaussian mixture model.

## I. INTRODUCTION

Histology is the science of understanding the structure of animals and plants, and studying the functional implications of biological structures. The knowledge of biological structures and their functions at the subcellular, cellular, tissue and organ levels is central to the understanding of mechanisms of disease and drug actions. Therefore, histology provides a scientific foundation for clinical research, education, and practice. Our previous study on human histological image classification [8] shows that boundary images account for a large portion of misclassification. Boundary images are those taken from tissues around the boundary of the slide or the boundary of the organ. They contain less or no conceptual information of the tissue/organ to be studied. Fig. 1 shows examples of boundary images that are frequently misclassified in our previous study. Training a model using boundary images will introduce irrelevant information, hence hampering the performance of the classifier. However, identifying sub-images that truly represent the tissue/organ requires domain knowledge. Manually extracting these sub-images from a large collection of images can be a very laborious task.

This kind of ambiguous and noisy labeling of training data is one example of the so called *Weak Labeling* problem. The Weak Labeling problem has attracted much attention. One active research area is called *multiple instance learning* (MIL) [5], in which training samples are given in the form of bags of instances; labels of bags are given, but labels of instances in bags are unknown. MIL has been applied in a variety of areas, such as drug activation prediction, DNA motif discovery, and data mining.

\*This work was supported in part by the Louisiana Board of Regents, The Research Institute for Children, and University of New Orleans.

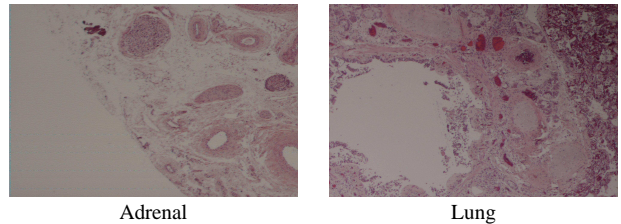


Fig. 1. Examples of boundary images that contain less or no conceptual information.

In this paper, histological image classification is tackled as a MIL problem. Each image is viewed as a bag with instances defined by subimages. We applied a novel learning method, MILES (Multiple-Instance Learning via Embedded instance Selection) [4], which converts the MIL problem to a supervised learning problem. 1-norm SVM is then applied to construct classifiers and select important features simultaneously.

## II. OUR APPROACH

### A. Image Feature Extraction

The images are first converted into LUV color space. Only the L component is used to extract features. This is because the color of histological images largely depends on the stained material and the L component is in a certain sense invariant to color variations. Histogram equalization is applied to the L component to eliminate luminance variances.

Histological images are essentially composed of texture patches. Textural images are usually distinguishable with scale and orientation features. Since texture patches in a histological image are in general inhomogenous, it's difficult to use global features to categorize histological images. Instead, a histological image is divided into sub-images. Multi-channel Gabor filters are used to extract texture features from sub-images.

A histological image is represented with a number of feature vectors each corresponding to one sub-image. A typical feature vector is represented as

$$[\mu_{0,0}, \sigma_{0,0}, \mu_{0,1}, \sigma_{0,1}, \dots, \mu_{S-1,O-1}, \sigma_{S-1,O-1}]^T$$

where,  $\mu_{sn}$  and  $\sigma_{sn}$  are mean and standard deviation calculated from Gabor responses corresponding to the filter with scale  $s$  and orientation  $n$ .

## B. MIL and MILES

Dietterich et al. [5] first formalized the MIL model and developed a MIL algorithm named axis parallel rectangles (APR) method. Maron and Lozano-Pérez [6] presented a framework, *Diverse Density(DD)* that extended the APR approach. Zhang and Goldman [7] developed an EM-DD algorithm that reduced the computational cost of the DD method. Andrews et al.[1] formulated MIL as a mixed integer quadratic program. Chen and Wang proposed a DD-SVM, which combines Diverse Density and Support Vector Machine (SVM) to deal with the problem that the underlying true concept may consist of several instance prototypes. Many standard supervised learning methods, such as,  $k$ -NN, neural network, and decision trees, have been adopted to solve the MIL problem.

The aforementioned algorithms either require a high computation cost, or are not capable of learning complex concepts. MILES tackles these limitations by converting a MIL problem to a standard supervised learning problem using instance-based feature mapping [4]. A feature is defined using each instance in a training bag. A bag is then mapped into a new feature space  $\mathcal{F}_c$ , the dimension of which is the total number of instances in all the training bags. Specifically, the embedding of bag  $B_i$  is achieved by the following mapping:

$$m(B_i) = [s(x^1, B_i), s(x^2, B_i), \dots, s(x^n, B_i)]^T$$

where  $x^k, k = 1, \dots, n$  are the instances from the training bags.  $s(x^k, B_i)$  is a measure of similarity between the instance  $x^k$  and the bag  $B_i$ , and is determined by the concept and the closest instance in the bag. The coordinates of a given bag in the feature space represent the bag's similarities to various instances in the training set.

The embedding produces a possibly high dimensional feature space when the number of instances in the training set is large. Many features may be redundant or irrelevant. MILES applies 1-norm SVM [2] to construct classifiers and select features simultaneously. Please refer to [4] for a detailed mathematical treatment of MILES.

## III. EXPERIMENT RESULTS

The algorithm is tested on 778 human histological images from 10 categories as shown in Table I. These images are of  $40\times$  magnification (4 objective lens  $\times$  10 ocular lens) stored in JPEG format with size  $3072 \times 3840$ . To reduce the computational cost, each image is downsampled to  $1536 \times 1920$ . The block size is  $64 \times 64$ . The Gabor filter bank consists of filters of 3 scales and 6 orientations.

Five fold cross-validation is conducted on these histological images. For each fold, 45 pairwise classifiers are trained using MILES. These 45 binary classifiers have very high classification rate, lots of them with 0 test error. This is because MILES can automatically identify sub-images that truly represent their organs, and use them to construct classifiers and to do classification. For example, most parts in the boundary images in Fig. 1 will be labeled as irrelevant, and excluded from classification. The experiment also shows good sparse characteristic of 1-norm SVM, e.g. for feature

TABLE I  
THE CATEGORY LABELS OF THE IMAGES AND 5-FOLD  
CROSSVALIDATION COMPARISON RESULTS.

| Histological Images |          |              | Test Accuracy Comparison |       |
|---------------------|----------|--------------|--------------------------|-------|
| Cid                 | Cname    | Image Number | Mixture Model            | MILES |
| C1                  | adrenal  | 25           | 0.48                     | 0.80  |
| C2                  | heart    | 114          | 0.808                    | 0.89  |
| C3                  | kidney   | 20           | 0.65                     | 0.75  |
| C4                  | liver    | 107          | 0.813                    | 0.92  |
| C5                  | lung     | 288          | 0.872                    | 0.88  |
| C6                  | pancreas | 120          | 0.608                    | 0.73  |
| C7                  | spleen   | 18           | 0.933                    | 0.75  |
| C8                  | testis   | 25           | 0.44                     | 0.60  |
| C9                  | thyroid  | 39           | 0.90                     | 0.84  |
| C10                 | uterus   | 22           | 0.73                     | 0.60  |
| Average             |          |              | 0.723                    | 0.776 |

space of dimension 1540, only 36 features were selected. These selected features are very representative, reflecting true concepts of two participating classes.

The multiclass classifier takes a majority vote on 45 pairwise classifier results and assigns the test bag to the class that wins. As shown in Table I, the average test accuracy is improved from 72.3% to 77.6%.

## IV. SUMMARY AND FUTURE WORK

We applied a MIL learning method, MILES, to human histological image classification. MILES converts a MIL problem to a supervised learning problem by instance-based mapping. 1-norm SVM is adopted to do feature selection and classification simultaneously. MILES demonstrates strong concept learning capability, thus can potentially be adapted to automated pathological image analysis, and other cause-effect biomedicine studies where causes are unknown.

## REFERENCES

- [1] S. Andrews, I. Tsochantaris, and T. Hofmann, "Support Vector Machines for Multiple-Instance Learning," *Advances in Neural Information Processing Systems 15*, pp. 561-568, 2003.
- [2] J. Bi, K. P. Bennett, M. Embrechts, C. Breneman and M. Song, "Dimensionality Reduction via Support Vector Machines," *Journal of Machine Learning Research*, vol. 3, pp. 1107-1135, 2003.
- [3] Y. Chen, J. Z. Wang, "Image Categorization by Learning and Reasoning Regions," *Journal of Machine Learning Research*, vol. 5, pp. 913-939, 2004.
- [4] Y. Chen, J. Bi and J. Z. Wang, "MILES: Multiple-Instance learning via Embedded Instance Selection," Under review for IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006. <http://www.cs.uno.edu/~yixin/date/pami06rev.pdf>
- [5] T. G. Dietterich, R. H. Lathrop and T. Lozano-Petez, "Solving the Multiple Instance Problem with Axis-Parallel Rectangles," *Artificial Intelligence*, vol. 89, nos. 1-2 pp. 31-71, 1997.
- [6] O. Maron and T. Lozano-Pérez, "A Framework for Multiple-Instance Learning," *Advances in Neural Information Processing Systems 10*, pp. 570-576, 1998.
- [7] Q. Zhang and S. A. Goldman, "EM-DD: An improved Multiple-Instance Learning Technique," *Advances in Neural Information Processing Systems 14*, pp. 1073-1080, 2002.
- [8] D. Zhao, Y. Chen, and H. Correa, "Statistical Categorization of Human Histological Images," *Proc. International Conference on Image Processing*, vol. III, pp. 628-631, Genova, Italy, 2005